PB-218 001

BASIC ENVIRONMENTAL STATISTICS NOTEBOOK

Environmental Protection Agency
Cincinnati, Ohio

March 1971

DISTRIBUTED BY:

NTIS
National Technical Information Service
U. S. DEPARTMENT OF COMMERCE
5285 Port Royal Road, Springfield Va. 22151

PB 218 001

Basic Environmental Statistics Notebook

CLEAN
WATER

Reproduced by
NATIONAL TECHNICAL
INFORMATION SERVICE
U S Department of Commerce
Springfield VA 22151

TRAINING MANUAL

WATER QUALITY OFFICE
ENVIRONMENTAL PROTECTION AGENCY

BASIC ENVIRONMENTAL STATISTICS NOTEBOOK

Conducted by
Special Training Branch
TRAINING INSTITUTE

Environmental Health Service
Environmental Control Administration

for

National Training Center
Water Quality Office
ENVIRONMENTAL PROTECTION AGENCY

MARCH 1971

214

FOREWORD

These manuals are prepared for reference use of students enrolled in scheduled training courses of the Water Quality Office, Environmental Protection Agency.

> *Due to the limited availability of the manuals, it is not appropriate to cite them as technical references in bibliographies or other forms of publication.*
>
> *References to products and manufacturers are for illustration only; such references do not imply product endorsement by the Water Quality Office, Environmental Protection Agency.*

The reference outlines in this manual have been selected and developed with a goal of providing the student with a fund of the best available current information pertinent to the subject matter of the course. Individual instructors may provide additional material to cover special aspects of their own presentations.

This manual will be useful to anyone who has need for information on the subjects covered. However, it should be understood that the manual will have its greatest value as an adjunct to classroom presentations. The inherent advantages of classroom presentation is in the give-and-take discussions and exchange of information between and among students and the instructional staff.

Constructive suggestions for improvement in the coverage, content, and format of the manual are solicited and will be given full consideration.

HM Freeman

H.M. Freeman
Chief, Direct Training Branch
Division of Manpower and Training
Water Quality Office
Environmental Protection Agency

TRAINING PROGRAM

Through the Water Quality Office, Environmental Protection Agency conducts programs of research, technical assistance, enforcement, and technical training for water pollution control.

Training is available at five installations of the Agency. These are: the National Training Center located at the Robert A. Taft Sanitary Engineering Center in Cincinnati, Ohio; the Robert S. Kerr Water Research Center, Ada, Oklahoma; the Southeast Water Laboratory, Athens, Georgia; the Pacific Northwest Water Laboratory, Corvallis, Oregon; and the Hudson-Delaware Basins Office, Edison, New Jersey.

The objectives of the Training Program are to provide specialized training in the field of water pollution control which will lead to rapid application of new research findings through updating of skills of technical and professional personnel, and to train new employees recruited from other professional or technical areas in the special skills required. Increasing attention is being given to development of special courses providing an overview of the nature, causes, prevention, and control of water pollution.

Scientists, engineers, and recognized authorities from other Agency programs, from other government agencies, universities, and industry supplement the training staff by serving as guest lecturers. Most training is conducted in the form of short-term courses of one or two weeks' duration. Subject matter includes selected practical features of plant operation and design, and water quality evaluation in field and laboratory. Specialized aspects and recent developments of sanitary engineering, chemistry, aquatic biology, microbiology, and field and laboratory techniques not generally available elsewhere, are included.

The primary role and responsibility of the States in the training of wastewater treatment plant operators are recognized. Technical support of operator-training programs of the States is available through technical consultations in the planning and development of operator-training courses. Guest appearances of instructors from the Environmental Protection Agency, and the loan of instructional materials such as lesson plans and visual training aids, may be obtained through special arrangement. These training aids, including reference training manuals, may be reproduced freely by the states for their own training programs. Special categories of training for personnel engaged in treatment plant operations may be developed and made available to the States for their own further production and presentation.

An annual Bulletin of Courses is prepared and distributed by the Water Quality Office. The Bulletin includes descriptions of courses, schedules, application blanks, and other appropriate information. Organizations and interested individuals not on the mailing list should request a copy from one of the training centers mentioned above.

ENVIRONMENTAL PROTECTION AGENCY
WATER QUALITY OFFICE
DIVISION OF MANPOWER AND TRAINING

R. F. Guay, Director

Training Grants Branch
Manpower Planning Branch
State and Local Manpower Development Branch
Direct Training Branch
Office of Environmental Activities

National Training Center
Robert A. Taft Sanitary
Engineering Center
4676 Columbia Parkway
Cincinnati, OH 45226

REGIONAL MANPOWER DEVELOPMENT AND TRAINING OFFICES

Northeast Region

Manpower and Training Office
John F. Kennedy Federal Building
Room 2303
Boston, MA 02203

Manpower and Training Office
Hudson-Delaware Basins Office
Edison, NJ 08817

Middle Atlantic Region

Manpower and Training Office
918 Emmett Street
Charlottesville, VA 22901

Southeast Region

Manpower and Training Office
Southeast Water Laboratory
College Station Road
Athens, GA 30601

Ohio Basin Region

Manpower and Training Office
Federal Office Building
Room 7027
Cincinnati, OH 45226

South Central Region

Manpower and Training Office
Robert S. Kerr Water Research Center
P. O. Box 1198
Ada, OK 74820

Missouri Basin Region

Manpower and Training Office
911 Walnut Street
Kansas City, MO 64106

Great Lakes Region

Manpower and Training Office
33 E. Congress Parkway
Chicago, IL 60605

Southwest Region

Manpower and Training Office
620 Central Avenue
Alameda, CA 94501

Pacific Northwest Region

Manpower and Training Office
Pacific Northwest Water Laboratory
200 South 35th Street
Corvallis, OR 97330

INTRODUCTION TO TESTS OF SIGNIFICANCE

the [illegible] the [illegible]

In this paper we characterize the program action necessary in a two-tailed test of significance by defining and relating the following: (1) test of significance, (2) statistic, (3) null hypothesis, (4) test assumptions, (5) law of random variation for the statistic when H_0 is true, (6) alternative hypothesis, (7) distribution of the statistic when H_A is true, (8) four possible decisions in a test of significance (9) type-I error, (10) critical values, or value, (11) type-II error, (12) power of the test, (13) rejection region, (14) acceptance region, (15) decision rules, (16) statistical decision, (17) diagram of a two-sided test of significance, (18) robustness, (19) sample size, (20) test of significance summarized, (21) example of a two-tailed t-test.

to p. 16

AN INTRODUCTION TO TESTS OF SIGNIFICANCE

Joseph F. Santner*

I INTRODUCTION

A Test of Significance

Statistics has been classified by many into two broad subject areas: descriptive statistics and statistical inference (see Figure 1).

Descriptive statistics makes extensive use of pictures, tables, graphs, and other visual arrangements. These methods are found popularized in magazines and newspapers. The onus is centered on summarizing the available information into an easily assimilated form and not on acquiring new knowledge about the world we live in.

Inferential statistics on the other hand is further subdivided into two classes of problems, namely, estimation and tests of significance (other names for the latter are hypothesis testing, significance tests, and tests of hypothesis). New knowledge is acquired via induction; that is we progress from the particular to the general or from sample to parent population.

Estimation methods are used to approximate numerical values of unknown population parameters (like the mean) from incomplete or sample data. A test of significance can be defined as a method of analyzing data so as to discriminate between two hypotheses.

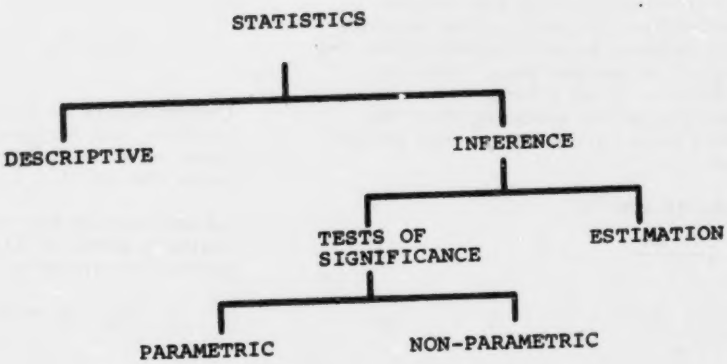
These two methods are closely related, and indeed one can contruct a procedure to cover both aspects. If they are separated, as done here, then the logical development for a test of significance is greatly simplified resulting in a comprehensive understanding in a shorter time period.

The test of significance in turn is only one small link' of the total experimental chain that can be briefly described as (1) problem definition, (2) data collection scheme, (3) data collection, (4) data analysis, and (5) the written report.

In the most carefully controlled investigation it is virtually impossible for two identical experiments to yield absolutely identical responses, under the assumption

FIGURE 1

CLASSIFICATION OF STATISTIC



---

*Mathematical Statistician, Special Training Branch, Training Institute, ECA.

that the measuring instrument is sensitive. If, however, it is not, then identical responses can always be obtained.

This inability to obtain identical responses is due to random variation in the surrounding environmental conditions, the imprecision of the measuring instruments, limitation of the observer's technique and experience, and the inherent variability encountered from sample to sample. All these factors plus others induce fluctuations in the experimental reponse. Even when the tested hypothesis is true the observed result will not match it.

The purpose then of a statistical test of hypothesis is to sort out and identify the differences attributable to expected random observational fluctuations as opposed to the differences attributable to deviations from the hypothesis.

The basic question to be answered is "How large must an observed difference be to justify rejecting the hypothesis?" Or stated otherwise, "How divergent must the observed difference be in order to be called a rare difference as opposed to expected random variation?"

To answer this question the procedure or the steps taken in a two-tailed test of hypothesis are always the same. The explanation of the variations on the basic two-tailed test will be taken up in the sequel.

The criterion or measure used for discriminating between differences or distinguishing between hypotheses is a statistic.

Statistic

A statistic (note no s on the end) can be simply defined as a function of the sample. That is, it is some value calculated from the observational or sample data. This is accomplished by using a formula or some prescribed recipe for mangling the data. The method varies from test to test and is not unique.

Two examples are

$$z = \frac{(\bar{x} - 10)\sqrt{n}}{5} \qquad (1)$$

$$t = \frac{(\bar{x} - 10)\sqrt{n}}{s} \qquad (2)$$

where $\bar{x}$ = sample mean
n = sample size
5 = standard deviation of parent population
s = sample estimate of standard deviation

In (1) the statistic value calculated from the sample is used to test if the mean of a sampled normal population is 10 when its standard deviation is known and equals 5. In (2) the test is the same except that the sampled population's standard deviation is unknown and is estimated from the data as s.

For example if the four sampled values are 10, 11, 12, and 11, then $z = (11 - 10)\sqrt{4} / 5 = 0.4$

The statistic is the measure or criterion used to discriminate between two hypotheses used in a test of significance, namely, the null hypothesis and the alternative hypothesis.

Null Hypothesis

The null hypothesis is defined as the hypothesis of no difference between a hypothetical and the sample population. It will be shown later that it should be formulated for the express purpose of being rejected or nullified (see Alternative Hypothesis).

If the hypothetical population is labeled as P_h and the sampled population as P_s, then "no difference" can be written as $P_s - P_h = 0$ or null. Many statisticians notationally write the null hypothesis as

$$H_O\colon\ P_s = P_h$$

or

$$H_O\colon\ P_s - P_h = 0$$

Unfortunately this notation is not universally uniform, but fortunately most deviations from these are centered in using a different subscript for H.

As an example the null hypotheses for the statistic given in (1) and (2) are both written the same as

$$H_O\colon\ \mu = 10.$$

For (1) the hypothetical population is normally distributed with mean = 10 and standard deviation = 5 while the sampled population is normally distributed with standard deviation = 5 and an unknown mean, and we wish to test whether the unknown mean = 10.

For (2) the hypothetical population is normally distributed with mean = 10 and an unknown standard deviation while the sampled population has unknown mean and standard deviation, and we wish to test whether the unknown mean = 10.

In both cases we test whether the difference between the sampled and hypothetical population is null or whether the sampled population has a mean of 10.

Restrictions such as the type of population sampled as well as parameter information have been placed. These restrictions are popularly called test assumptions.

Test Assumptions

Each test of significance has associated with it explicit or implicit, or both, test assumptions that must be satisfied if the results of the test are to be valid. Failure to meet the assumptions affects the probability statements that will be introduced later under the topic Type-I Error (see also Robustness).

For example the test assumptions or the mathematical model for (1) are that the sampled population is normally distributed with standard deviation = 5. The theory also requires that the samples be random and independent.

Test assumptions are generally classified as either parametric or non parametric. The former have more stringent requirements, and justification for their use is generally more difficult. For example that the sampled population is normally distributed with a known standard deviation could be a parametric test assumption.

A non parametric test requires no hypothesis about specific parameter values. These tests are often called distribution-free, which really means that the test is independent of the form of the underlying distribution. These two are not interchangeable but are so often done that usage has made them indistinguishable. A non parametric test might merely require that the sampled population be continuous in order to be validly applied. The requirements for a non parametric test are generally fewer and can be applied to a larger set of source data, or used when the parametric assumptions cannot be proved or are unknown.

The advantages of parametric exceed non parametric when applicable and will be discussed further under the topic Power of the Test.

Let us assume that the test assumptions (or assumption) are (or is) satisfied. If the null hypothesis is true (sampled and hypothetical populations are equal) then the sampled values are determined by the sampled population. These sampled values fix the value for the statistic. Different random samples result indfferent random values for the statistic; that is we say the statistic is a random variable. The law of random variation that the statistic follows when the null hypothesis is true must be known before discrimination is possible.

Law of Random Variation for the Statistic When H_0 is True

The law of random variation for the statistic can be determined either analytically or empirically. The analytic determination is made by the originator of the test, who is a researcher or a mathematical statistician.

For specific examples of an analytic development the reader is referred to the journals or the many commercial texts available on mathematical statistics.

Some examples of the statements of several of these laws as determined by the researcher might be (1) the statistic is normally distributed with zero mean and unit variance, or (2) the statistic is distributed as chi-square with degrees of freedom equal to sample size minus one, or (3) the statistic is t-distributed with degrees of freedom equal to sample size minus two.

The empirical determination of the law of random variation can be done by anyone who really understands the test that is being applied. If the experimenter cannot imagine some sampling sequence using random numbers to determine the law of variation then there is a deficiency somewhere.

As an example the empirical determination for the z statistic defined by (1) could be as follows. Take a random sample of nine, $n = 9$ (any value for n can be used), from a normal population whose mean is 10 and whose standard deviation is 5. From these nine values calculate $\bar{x}$ and finally the value of z from equation (1). As a result of this first sequence we have one value for z; label it z_1.

This sequential procedure can be repeated as often as desired. Suppose it is repeated 1,000 times, resulting in 1,000 different values for z, namely $z_1, z_2, \ldots, z_{1,000}$. We can form a histogram using these 1,000 values and then draw a smooth, continuous curve through the center of the rectangles as shown in Figure 2.

The smooth curve labeled H_0 in Figure 2 is an approximation to the law of random variation or the distribution of the z statistic when H_0 is true. If instead of using 1,000 values for z, we used all possible values then the distribution would have been exact, that is, in perfect agreement with the analytic solution. With 1,000 values the differences between the two solutions are small and furthermore can be made arbitrarily small by calculating enough sequences. The approximate solution converges or approaches the analytic solution as the number of sequences increases.

If the above experiment were performed then it would be determined that z would be approximately normally distributed with zero mean and unit variance, regardless of the value of n used.

The t statistic in (2) when sampling from normal is distributed as a t-distribution with degrees of freedom equal to $(n - 1)$ provided that the degrees of freedom used in estimating s is $(n - 1)$. Usually this is true since $\bar{x}$ and s are estimated from the same data.

That these distributions do approximate the ones claimed can be verified objectively via a test of significance or subjectively by using the eyeball test. With 1,000 samples the agreement should pass the eyeball test easily by the comparing of percentiles or some other cumulative values.

The reader is urged to verify (1) when $n = 4$ or some small number. If a computer is used then 1,000 is a small number of sequences. If a computer is not used then decrease the number of sequences, realizing that the approximation will not conform as well.

There are four distributions that one can talk about up to this point: (1) the hypothetical population or distribution, (2) the sampled population, (3) the distribution of the random sample of size n taken from the sampled distribution, and (4) the distribution or law of random variation for the statistic calculated from the sample (see Figure 3A).

FIGURE 2

DISTRIBUTION OF THE STATISTIC WHEN THE NULL HYPOTHESIS IS TRUE

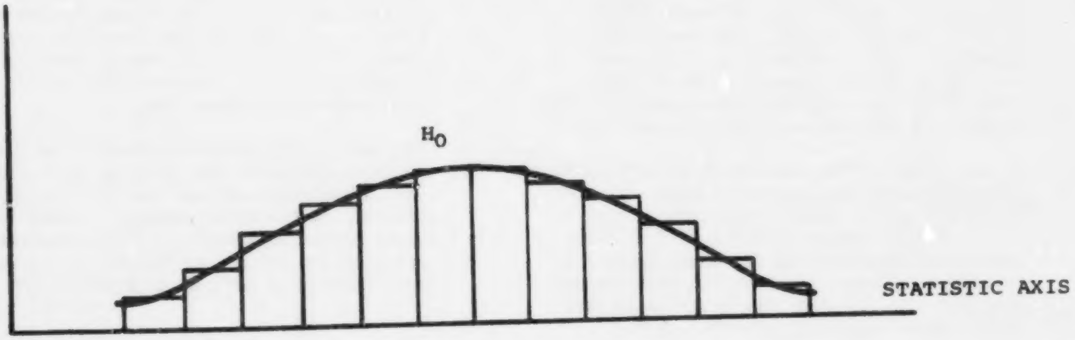
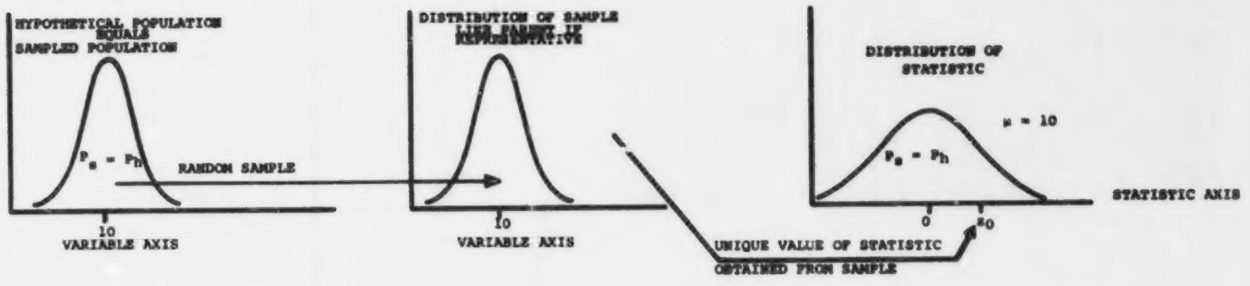


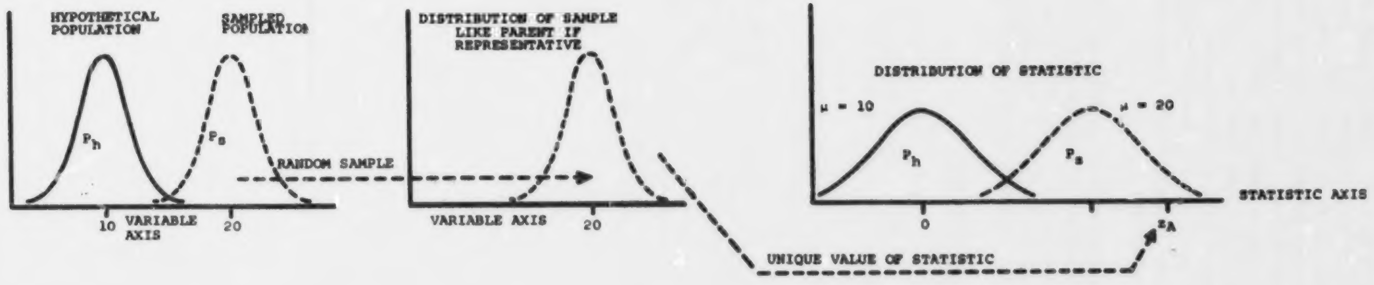
FIGURE 3

FOUR DISTRIBUTIONS IN HYPOTHESIS TESTING

H_0 IS TRUE


HYPOTHETICAL POPULATION EQUALS SAMPLED POPULATION
P_s = P_h
RANDOM SAMPLE
10
VARIABLE AXIS
DISTRIBUTION OF SAMPLE LIKE PARENT IF REPRESENTATIVE
10
VARIABLE AXIS
DISTRIBUTION OF STATISTIC
μ = 10
P_s = P_h
STATISTIC AXIS
0
z_0
UNIQUE VALUE OF STATISTIC OBTAINED FROM SAMPLE

H_A IS TRUE


HYPOTHETICAL POPULATION
SAMPLED POPULATION
P_h
P_s
RANDOM SAMPLE
10
VARIABLE AXIS
20
DISTRIBUTION OF SAMPLE LIKE PARENT IF REPRESENTATIVE
VARIABLE AXIS
20
DISTRIBUTION OF STATISTIC
μ = 10
μ = 20
P_h
P_s
STATISTIC AXIS
0
z_A
UNIQUE VALUE OF STATISTIC

In an actual test there is only one sample and hence only one value for the statistic. In Figure 3A it is labeled z_0. The reader should bear in mind that it is only one value out of a possible range of expected random variation. Some intervals, especially those near the mean, have higher probabilities than others, owing to the nature of the distribution. It will be shown later that these probabilities will be used to determine unexpected random variation.

If H_0 is not true then in the real world some alternative hypothesis is true, and this circumstance gives rise to a different law of random variation, owing to the truth of the alternative hypothesis.

Alternative Hypothesis

If the null hypothesis is not true then the negation of this is the alternative hypothesis. For our example earlier this is written by most as

$$H_A: \mu \neq 10$$

and read as "the mean of the sample population is not equal to 10." Again observe the caution that this notation varies from author to author. The form of the test dictated by such an alternative hypothesis is a two-sided or two-tailed test, and the reason is shown in Figure 4. There it will be noted that the alternative hypothesis occupies both sides of the parameter space, that is, to the right and left of $\mu = 10$. A one-sided or one-tailed test will be discussed in the sequel, and as the reader has guessed, the alternative occupies one side of the parameter space.

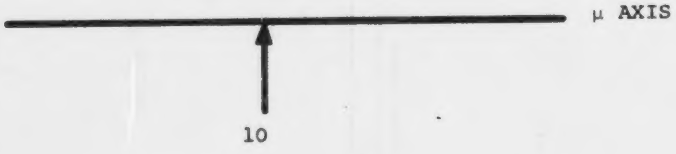
The alternative hypothesis must be the operational statement of the experimenter's research hypothesis. If this is not so, or the null hypothesis is the research hypothesis, then the researcher cannot really prove his research hypothesis. In order to prove the research or alternative hypothesis the null hypothesis must be rejected or nullified. A heuristic argument justifying this very important position (alternative and research hypothesis should be the same) will be taken up under Decision Rules.

In the design of experiments, data collection, and the subsequent test of hypothesis, the aim of the researcher should be to reject his favorite theory.

If the result of the tests of significance from data so collected does reject the pet theory then savings in money, manhours, and material have been realized. On the other hand if the test results do not reject his position the implication is a large gain in the confidence of the validity of the theory. One is in the

FIGURE 4

PARAMETRIC SPACE FOR A TWO-SIDED TEST



$H_0 : \mu = 10$

$H_A : \mu \neq 10$

position of trying to prove something is wrong but being unable to do so. Each failure adds extra strength to the belief that it cannot be wrong.

Too often researcher design data collection to support their theory and innocently build in unknown biases. By taking the opposite position and vigorously attacking it the danger of bias is decreased. An aggressive philosophy will yield more rapid advances in research.

Distribution of the Statistic When H_A is True

Figure 3A shows the distribution of the statistic z in (1) when H_0 is true or the sampled population has $\mu = 10$. How is that distribution affected if the alternative hypothesis, $\mu \neq 10$ is true?

To be specific, suppose the sampled population had $\mu = 20$ (see Figure 3B). By examining the numerator of (1) we note that z would now tend to have mostly positive values since most sample means would be greater than 10. This distribution for z would be shifted to the right of the $\mu = 10$ curve.

A sequential sampling scheme would give its approximate shape. Obviously it would be some distribution like that labeled $\mu = 20$ in Figure 3B. If $\mu = 0$ then by the same argument the distribution for the statistic would be shifted to the left of the $\mu = 10$ curve, as shown in Figure 5.

By applying the sequential sampling process it is easily seen that, for every possible value of μ for the sampled population, there would exist some distribution for the statistic. It could be any one of the two shown in Figure 5 or some other not drawn there.

Regardless of the value of μ in the sampled population the reader can imagine one but only one distribution for H_A if true. If H_0 is true then there is only one distribution for the statistic (see Figure 3A).

The uncertainty faced is what is the position for the true sampled population? Is it coincident as in 3A or shifted as in 3B?

This problem is solved by partitioning the statistic axis so as to maximize our possible correct decisions and minimize our possible incorrect decisions when we discriminate.

Four Possible Decisions in a Test of Significance

The correct decisions are (1) do not reject H_0 when it is true and (2) reject H_0 when it is false. Our incorrect decisions are (3) reject H_0 when it is true, or (4) do not reject H_0 when it is false. These four possible decisions are shown in Figure 6. In order to distinguish between the two types of incorrect decisions or errors, they are given the special names type-I and type-II error.

FIGURE 5

SOME POSSIBLE DISTRIBUTIONS FOR THE STATISTIC

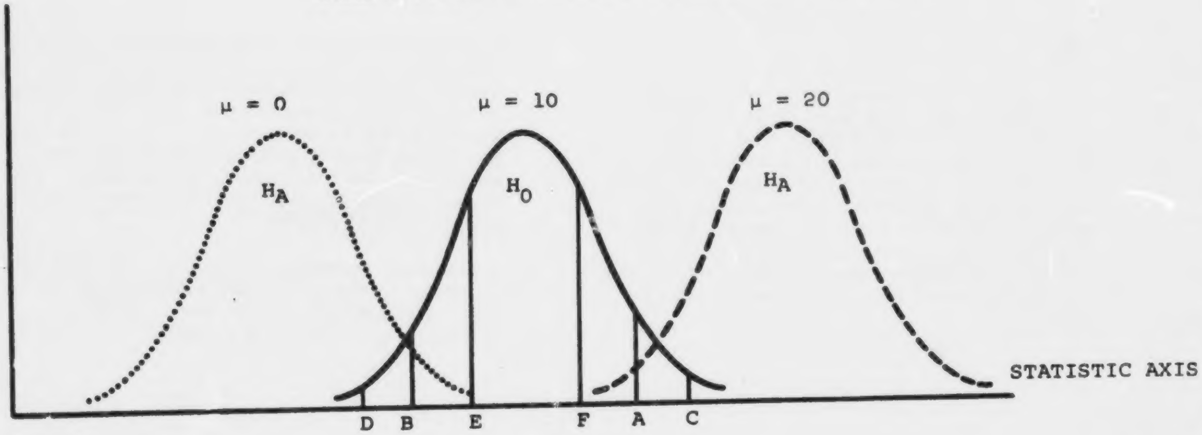


FIGURE 6

FOUR POSSIBLE DECISIONS IN A TEST OF SIGNIFICANCE

		STATISTICAL DECISION	
		DO NOT REJECT H_0	REJECT H_0
REAL	H_0 IS TRUE	CORRECT DECISION	INCORRECT DECISION (TYPE I ERROR) (α-ERROR)
WORLD	H_0 IS NOT TRUE	INCORRECT DECISION (TYPE II ERROR) (β-ERROR)	CORRECT DECISION

Type-I Error

The type-I error (also called α - error or error of the first kind) is the rejection of the null hypothesis when it is actually true. The seriousness or quantification of this error is measured by using probability, and its value is obtained in the following way. Suppose in Figure 5 that we adopt the following discrimination rule. If the statistic is greater than or equal to A(some constant) reject H_0; otherwise do not reject H_0.

Suppose the sampled population is $\mu = 10$. Since the distribution of the statistic for the hypothetical population is always known, its area under the curve to the right of a known A can be determined. This area is the probability of getting a value for the statistic, which according to our rule means reject H_0 falsely, or a type-I error has been made.

A familiar example of making a deliberate α- error is the fairy tale wherein the little girl jokingly cries "wolf, wolf" falsely, H_0: no wolf, H_A: wolf.

The magnitude of the α-error is called α and is also known as the level of significance. In a test of significance the reverse procedure is actually followed, that is, α is pre-specified before any data are collected, and from a set of appropriate tables the corresponding value for A is determined.

It should be clear from Figure 3 that the law of random variation is determined by the test assumptions, and failure to meet them results in a different law. Using the incorrect law implies an incorrect A. Thus, conforming to the test assumptions is important for a correct A and α. The only loophole possible is discussed under robustness.

The tables for various laws of random variation are found in the back of almost every book in statistics and are used to specify A, which is called the critical value or significant limit.

Critical Values or Value

In the two-sided tests used here two critical values are used in order to maximize both correct decisions shown in Figure 6. A heuristic argument leading to this conclusion is as follows.

Assume that H_0 is true; therefore the only distribution on Figure 5 is the $\mu = 10$ curve. Let the statistic be the normally distributed z of (1). If $\alpha = 0.05$ then from a set of normal-curve tables the single critical value for the statistic is $A = 1.64$. The probability of z less than or equal to 1.64 is 0.95, or the correct decision (do not reject H_0) will be made, on the average, 95 percent of the time. The probability of getting z greater than 1.64 is 0.05. We shall make the incorrect decision or an α-error (reject H_0 when true), on the average, 5 percent of the time. The value for A is determined by α, and the distribution of the statistic is determined by the known hypothetical distribution whether H_0 is true or not.

Is A a good choice when H_0 is not true? Let us assume for definiteness that $\mu = 0$ in the sampled population (see Figure 5). The probability of a correct decision now (reject H_0) is almost zero. The value of the statistic would be some point on the $\mu = 0$ curve as determined by the particular sampled data. The probability that it is greater than A is the area under $\mu = 0$ to the right of A or almost zero. The correct decision of rejecting H_0 is almost never a poor choice for A under the assumption $\mu = 0$.

On the other hand A is a good choice if the sampled population is $\mu = 20$. In that case the probability of rejecting H_0 is high and is the area under the $\mu = 20$ to the right of A. If one alternative is true we have high probability of making the correct decision (reject H_0) while if another alternative is true we have low probability.

We try to correct this situation by the following procedure. Keep $\alpha = 0.05$ and change the critical value from A to $B = -1.64$. The decision rule is changed to reject H_0 if the statistic is less than B; otherwise do not reject H_0. Applying the same logic as in the previous two paragraphs shows that if $\mu = 20$ is true then the correct decision (reject H_0) has almost zero probability. If $\mu = 0$ is true then the correct decision has high probability.

Thus to protect against the alternative's being on either side of $\mu = 10$, the area should be put equally in both tails, giving rise to two critical values. For example, if $\alpha = 0.05$ then $A = 1.96$ and $B = -1.96$ for the normal curve. Do not reject H_0 will be the decision if the statistic lies between these values; otherwise it will be rejected.

Values of the statistic from 1.96 to -1.96 are expected random variation. Those outside this range are unexpected and represent genuine differences when the power of the test is considered. A topic to be defined soon.

If the sampled population is $\mu = 20$ then the probability of a correct decision (reject H_0) is the area under its curve to the right of A and to the left of B, a nice high number.

If the sampled population is $\mu = 0$ then the probability of a correct decision is the area under its curve to the right of A and to the left of B, also a high value.

If the sampled population is $\mu = 10$ then the probability of a correct decision (do not reject H_0) is the area under its curve between A and B, which $1 - \alpha$, a preselected number.

Both correct decisions have a high probability regardless of the value for μ in the sampled population.

The α is preselected by the researcher, not the consulting statistician; however, the implications of a large or small α should be made clear.

In initial research an α-error or 10 percent or even 20 percent is reasonable and not unusual. In that case the researcher is guarding against rejecting the research hypothesis and is less concerned about an α-error. Large α means that the distance between A and B in Figure 5 is decreased to, say, E and F. The decreased distance means that the continue research or reject H_0 probabilities are increased regardless of the sampled population.

As research continues and other tests are made, the α level is gradually reduced. If an α-error (to incorrectly continue research) was made initially, a subsequent test will reveal the mistake, and no harm will have been done except extra research.

On the other hand if the α-error is made small initially, the distance between A and B is increased to, say, C and D. The increased distance means that the do not reject H_0 (abandon research) probabilities are increased regardless of the sampled population.

If a type-II error (see Figure 6; abandon research when it should be continued) is made then the mistake would never be uncovered unless another researcher works in the same area.

A type-I error can also be thought of as the researcher's risk of following a false clue in research. If a type-I error is made then the statistical decision is that the research hypothesis is true when it really is not. The result is further research in a false area. Eventually enough evidence or subsequent tests of hypothesis will reverse the decision and reject the research hypothesis. Because of the initial incorrect decision, we have followed the false clue (continue research) and did subsequent research until a later test corrected the initial incorrect decision.

The type-I error can also be viewed as taking a risk. The risk is the probability of saying the research hypothesis is true when it really is not. This risk is α, and $1 - \alpha$ is the probability or the confidence one would like for accepting a true H_0.

Both values are related to each other by odds. The odds of being right when H_0 is true are $1 - \alpha$ to α. Therefore, if one chooses $\alpha = 0.05$ the odds of a right decision when H_0 is true are 0.95 to 0.05 or 19 to 1. Changing α to 0.01 changes the odds to 99 to 1, and $\alpha = 0.1$ implies odds of 9 to 1 of correct decision if H_0 is true.

One cannot set the odds too high since there is a penalty for each increase, namely, raising the probability of a type-II error. These two errors are not independent, and the interrelationship between them can be explained in the following way.

Type-II Error

The type-II error (also called β-error or error of the second kind) is the probability of not rejecting H_0 when it is false or is the probability of rejecting H_A when it is true. The β is the magnitude of the β-error.

Graphically this error can be shown in Figure 5 by considering the following hypothetical problem. Test H_0 with the decision do not reject H_0 if the statistic lies between A and B; otherwise reject. Suppose H_A is true and $\mu = 20$. The probability that a statistic will lie between A and B is the area under its curve between these two points. Therefore β, the probability of not rejecting H_0 falsely (it should be since H_A is true) is the area under the appropriate true alternative hypothesis over the region of do not reject H_0, or from A to B.

If the critical values are changed from A, B to C, D (decrease α) then β would be increased. Increasing α from A, B to E, F would decrease β. Thus α and β are inversely related.

Unfortunately the value for the alternative is never known and β cannot be calculated. The size of the type-II error depends on the disparity or distance between the null and alternative hypotheses, sample size, α, and the particular test of hypothesis used. Further remarks will be made about β estimation in the sequel.

The α-error is generally considered by most to be more serious than a β-error. Making false claims about our research, α-error, is more important error than abandoning a possibly fruitful research effort, β-error.

Furthermore, journals frequently publish articles in which H_0 is rejected, and one seldom sees articles that do not reject H_0. This mean α-errors are the only kind that can appear in print. One wonders how much duplication there is in research because of that fact. Failure to reject H_0 is also information that could guide other researchers in the selection of their hypothesis.

A good example of minimizing α irrespective of the effect on the size of β exists in our courts of law (H_0: not guilty, H_A: guilty). We would sooner accept a β-error (let a guilty man go free) than commit an α-error (send an innocent man to jail). Of course, we can guarantee $\alpha = 0$ always by setting everyone free.

The β-error can be thought of as the risk of failing to following a true clue. Our decision is do not reject H_0 (or reject research hypothesis) when H_0 should be rejected (or H_A accepted). Owing to this incorrect decision, we abandon further research that should be pursued. The true clue is the truth of H_A, but because of a β-error we reject it and abandon follow-up research.

Power of the Test

One minus beta $(1 - \beta)$ is defined as the power of the test. It is the probability of accepting H_A when true. Graphically it is the area under the true alternative over the rejection region.

It can be thought of as our confidence level of accepting the research hypothesis when true. Thus tests with high power are desirable. The odds of accepting H_A if true are $(1 - \beta)$ to (β).

If you have a set of data for which there is a parametric and a non-parametric test available then the parametric is more powerful and preferred. Suppose you were searching for someone in a large crowd, and the only clue you had was that the individual was definitely in the crowd. This search would not have much hope of being fruitful (i.e., powerful). On the other hand if the available clues were that the person sought is male, 6 feet tall, has black hair, is wearing a brown suit, smokes cigars, limps, has an amputated right arm, and so on, then the search would be effective or powerful. More clues imply a more powerful search, and by analogy the more assumptions that are made and realized, the more powerful the test.

Rejection Region

The preselection of α in conjunction with the known distribution of the statistic for P_h determines the critical values. These critical values in turn partition the total statistic axis into two parts, the rejection region (also called significance region or critical region) and the acceptance region.

The rejection region is so selected that if H_A is true then the probability of accepting it is high $(1 - \beta)$ while at the same time if H_0 is true the probability of rejecting it is α, the preselected value.

These values are classified as real differences, not random vairation, and represent a compromise when type-I and type-II errors are taken into consideration.

In Figure 5 if C and D are the critical values then the rejection region is to the right of C and to the left of D. The remainder of the statistic axis is called the acceptance region.

Acceptance Region

The acceptance region is selected so that if H_A is true the probability of rejecting it is small (β) while at the same time if H_0 is true the probability of not rejecting it is the preselected $1 - \alpha$.

The action to be taken when a statistic is in either region is called the decision rule.

Decision Rules

Our first decision or discrimination rule (see Figure 6) is do not reject H_0 if the statistic falls in the acceptance region (some aliases: cannot reject H_0, a nonsignificant result, outside the rejection region, outside the critical region, accept H_0, research hypothesis not proved).

A nonsignificant result should be interpreted as meaning that such a statistic value from these sample sizes this large is obtained so frequently when H_0 is true that the data convinces no one that anything more than a H_0 random process produced them. If H_0 is true then the acceptable odds, preselected by the researcher at $1 - \alpha$ to α, favor this decision. It also means that the research hypothesis was not proved statistically according to the standard of evidence based on probability.

The contracted statement of the decision rule, to accept H_0, is often used instead of the longer statement. If one took this statement literally (unfortunately some do) it would be incorrect, since the null hypothesis is never proved when a nonsignificant result is obtained.

A heuristic argument taken from Draper and Smith(1) is as follows. John Doe, an office worker, is at lunch when the idea is advanced that he is not rich, H_0: not rich, H_A: rich. Two proofs are cited: (1) he always buys his clothes at second-hand stores (2) he always brings his lunch or eats at the cheapest places as today.

The statistical parallel is that two nonsignificant tests of significance were incorrectly interpreted as proving H_0 as true instead of do not reject H_0. It was learned later that Mr. Doe died and left $500,000 to charity. All the proofs of the null hypothesis are invalidated.

One can never prove the null hypothesis by getting a nonsignificant result. The implication is clear: given two tests (1) H_0 = research hypothesis and (2) H_A = research hypothesis then one must choose the second test. If a significant result is obtained then the research hypothesis has been proved. See the sequel for further remarks.

Our second decision rule is reject H_0 if the statistic falls in the rejection region. Rejection means that if H_0 is true then a statistic value of this magnitude from a sample size this large is so rarely obtained (odds favoring it are α to $1 - \alpha$) by a random process alone that this peculiar statistic value points to something over and above the H_0 random process.

When the statistic falls within the rejection region, two logical possibilities exist. The first is that H_0 is true and a rare event or an α error by the random process. The only other possibility is that H_A is true and caused this large statistic value. Since the acceptable α-error was preselected the decision must be that the sampling was from H_A.

Some authors state the result of a significant statistic as H_A is true unless a rare event has happened. Constant repetition of this long statement becomes tedious and it is often contracted as H_A is true, accept research hypothesis, reject H_0, there is a significant result, and so forth.

Statistical Decision

The statistical decisions in a test are (1) do not reject H_0 if the statistic is in the acceptance region (2) reject H_0 if the statistic is in the rejection region.

In a test where the statistic is a continuous distribution the critical value can be in either region since it does not affect the value of α.

In tests where the statistic is discrete the tables clearly state whether the critical value is the boundary of the rejection or of the acceptance region. Caution is advised.

Note also that for discrete distributions probabilities vary by discrete jumps. Hence it is unusual to have α equal exactly to 0.05 or any other level. In those cases the α level is specified as 0.05 or less.

Diagram of a Two-Sided Test of Significance

The proceeding terms have been collected and their interrelationship shown in Figure 7. Some interpretations are as follows.

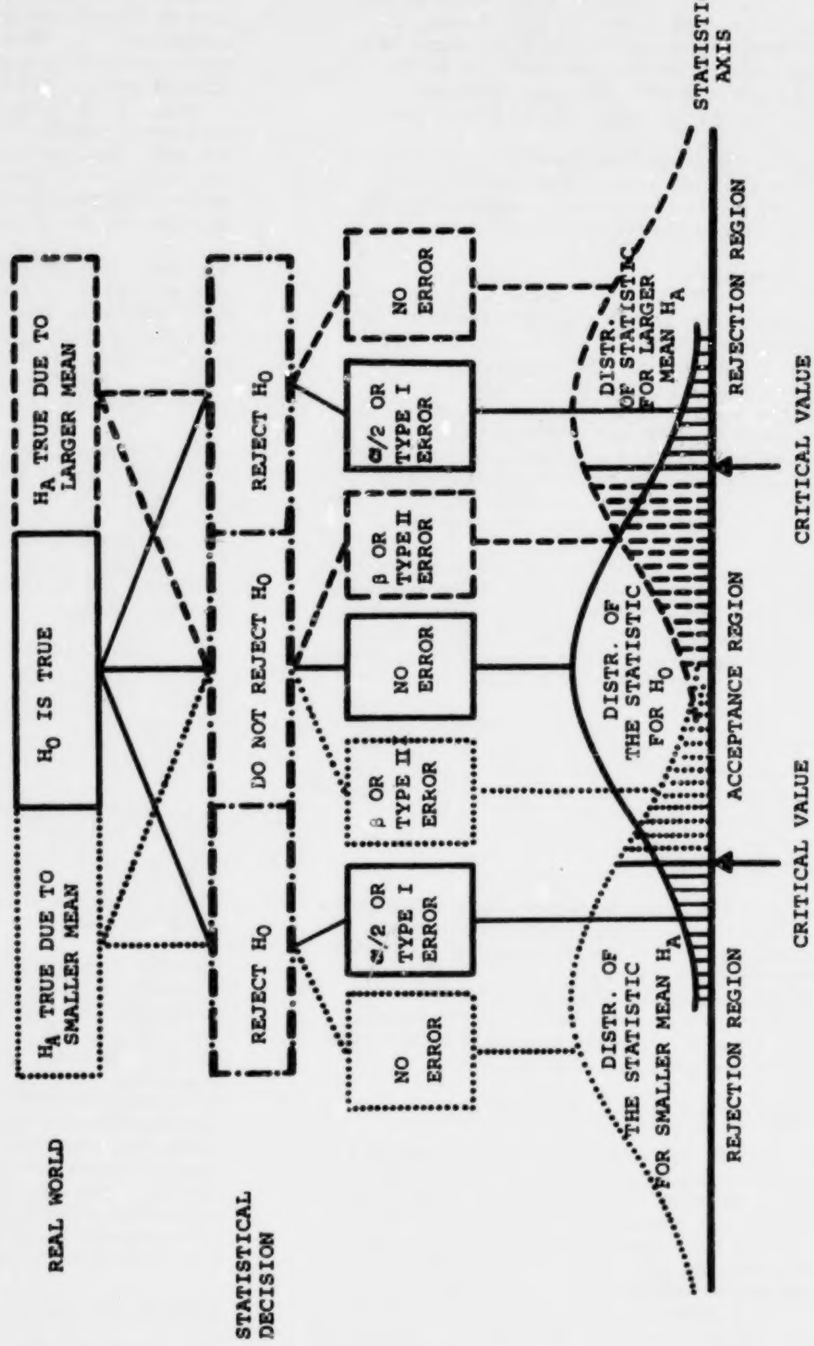
If H_0 is true or it is the only distribution for the statistic in Figure 7 then three kinds of statistic values can be obtained.

First would be a value in the left rejection region meaning reject H_0 incorrectly or an α-error. Next a value in the right rejection region with the same result. The third value we can obtain lies between the two critical values meaning do not reject H_0 correctly or no error.

On the other hand if the sampled population has a larger mean then P_h, its distribution, is on the right. When the first of its three values is in the right rejection region then H_0 is rejected or H_A is accepted correctly

FIGURE 7

TWO-SIDED TEST OF SIGNIFICANCE


REAL WORLD
H_A TRUE DUE TO SMALLER MEAN
H_0 IS TRUE
H_A TRUE DUE TO LARGER MEAN
STATISTICAL DECISION
REJECT H_0
DO NOT REJECT H_0
REJECT H_0
NO ERROR
$\alpha/2$ OR TYPE I ERROR
β OR TYPE II ERROR
NO ERROR
β OR TYPE II ERROR
$\alpha/2$ OR TYPE I ERROR
NO ERROR
DISTR. OF THE STATISTIC FOR SMALLER MEAN H_A
DISTR. OF THE STATISTIC FOR H_0
DISTR. OF STATISTIC FOR LARGER MEAN H_A
STATISTIC AXIS
REJECTION REGION
ACCEPTANCE REGION
REJECTION REGION
CRITICAL VALUE
CRITICAL VALUE

or no error. The next value could be in the left rejection region (not shown, to prevent overcrowding) with the same result. The third number could be in the acceptance region with the incorrect decision do not reject H_0 or a β-error.

The discussion if the sampled population has a smaller mean follows analogously and is left to the reader. Again only two of the three values are shown.

Robustness

A robust test is one that is insensitive to departures from the assumptions. A test that is sensitive to departures from the test assumptions lacks robustness. If a test is robust and the assumptions are not violated badly then the odds still remain as specified. This means that, for example, if one uses a test which assumes normality and the source data are quasi-normal, then a robust test can be confidently applied.

It is impossible to quantitate the departures from test assumptions that can be made in general and still have valid odds. Each test is different and much research has been done. The reader should consult with a statistician on these matters.

If a test requires several assumptions then one can talk about robustness with respect to each assumption or with respect to any combination of assumptions. See the sequel for further remarks.

Sample Size

A realistic question before data collection is how many sample points are needed. In many cases this question can be answered provided the researcher can supply information about variability and the size of differences that are to be discriminated. Consultation is again suggested since much work has been and is being done.

Test of Significance Summarized

A test of significance can be defined as a method of analyzing data so as to discriminate between two hypotheses. The first is the null hypothesis; the second is the alternative hypothesis, which should be the operational statement of the experimenter's research hypothesis. Next, the best test for the alternative hypothesis is selected, that is, the one with the greatest power, and is robust. This test is based upon assumptions, which may be explicitly or implicitly stated, or both. The probability statements used in discrimination are based upon the assumptions. After the test is selected, the α-error is prespecified, that is, the odds of being right if the null hypothesis is true. This α- error in conjunction with the known distribution of the statistic for the null hypothesis determines critical values. The critical values partition the entire statistic axis into two parts, the acceptance and rejection region.

The sample size is fixed and the data are now collected. From the collected data the statistic is evaluated and compared with the critical values to determine whether it is in the acceptance or rejection region. After this comparison the appropriate decision is made.

If the statistic is in the acceptance region, the appropriate decision is do not reject H_0, or cannot reject H_0, or reject H_A, and so on. If the statistic is in the rejection region, the decision is reject H_0, or accept H_A, or there is a significant difference, and so forth. With each decision there is a possible error.

When the statistic falls in the acceptance region and if in the real world H_A is true and we reject H_A, this is a type-II error; however, if H_0 is true there is no error. There is no chance for a type-I error when the statistic is in the acceptance region.

When the statistic falls in the rejection region and if in the real world H_0 is true and we reject H_0, this is a type-I error; however if H_A is true there is no error. There is no chance for a type-II error when the statistic is in the rejection region.

Briefly when a test of significance is performed the following information should be available: (1) H_0, (2) H_A, (3) α, (4) test assumptions, (5) statistic, (6) distribution for the statistic, (7) critical values, (8) acceptance region, (9) rejection region, and (10) decision after the statistic is evaluated. Additional desirable information is (11) a comparison of the power of the test versus that of others available for the same H_0 and H_A, (12) sample size needed to discriminate differences of specified magnitudes, (13) information about robustness, and (14) type-II error information.

Example of a Two-Tailed t-Test

We illustrate a two-tailed t-test by using a four-step process for convenience's sake.

Assume that we wish to test at the 5 percent significance level whether a new product average μ is different from the previous standard $= 7.35$. The observational data are random independent normally distributed X_i $i = 1, 2, \ldots, 25$.

Step 1. H_0: $\mu = 7.35$

H_A: $\mu \neq 7.35$

$\alpha = 0.05$

Step 2. Statistic $= \sqrt{n}\,(\bar{X} - 7.35)\,/\,s$, which is distributed as t-distribution with $24 =$ degrees of freedom. From a table of t values, the two critical values are $\pm\ 2.064$. The acceptance region lies between these two values, and the rejection region is all other statistic values.

Step 3. From the data $\bar{X} = 7.1$, $s = 0.504$; therefore, the statistic $= \sqrt{25}\,(7.1 - 7.35)/0.504 = -2.480$, or in the left rejection region.

Step 4. Reject H_0 since the statistic is in the rejection region, or the new average is different from the old one.

Many more examples can be found in almost every textbook on statistics. The four-step process used above is not generally used but is adopted here for instructional convenience.

In the concluding article we discuss one-tailed test, variations on a two-tailed test and conclude with several remarks that should clarify further other ideas about a test of significance.

REFERENCE

1 Draper, N.R. and Smith, H. Applied Regression Analysis, 1967, John Wiley and Sons, Inc., New York.

VARIATIONS ON A TEST OF SIGNIFICANCE

~~This is~~ the second ~~of a~~ two-part paper which describes the program of action in a test of significance when extended to one-sided tests. In addition simple and composite hypotheses are illustrated by several examples. Finally, the relationship between a significance test and confidence limits is exhibited when variations on a two-sided t-test are discussed.

VARIATIONS ON A TEST OF SIGNIFICANCE

Joseph F. Santner

In part one of this paper we characterized the program of action necessary in a two-tailed test of significance by defining and relating the following: (1) test of significance, (2) statistic, (3) null hypothesis, (4) test assumptions, (5) law of random variation for the statistic when H_0 is true, (6) alternative hypothesis, (7) distribution of the statistic when H_A is true, (8) four possible decisions in a test of significance, (9) type-I error, (10) critical values or value, (11) type-II error, (12) power of a test, (13) rejection region, (14) acceptance region, (15) decision rule, (16) statistical decision, (17) diagram of a two sided test of the significance, (18) robustness, (19) sample size, (20) test of significance summarized and (21) example of a two-tailed t-test. The purpose of this paper is to extend the procedure to one-sided tests.

I TWO-SIDED VERSUS A ONE-SIDED TEST

The decision to make either a two-sided test or a one-sided test (also called one-tailed) is not left to the discretion of the researcher. This determination is made in advance of any data collection and is based upon either theoretical grounds or previous experience. If at least one of these causes is present then a one-sided test can be made and is actually preferred since it is generally more powerful than the corresponding two-sided test. If, however, there exists no valid justification for a one-sided test, a two-sided test is mandatory. This procedure as well as the preselection of α before data collection guarantees that statistical decisions will not be biased.

II STATEMENT OF THE HYPOTHESES

The null and alternative hypotheses for a one-sided test can be written in two different ways, neither of which is universally acceptable to statisticians and neither of which is both logical and satisfying to the researcher. Authors recognize the problem; however, their varied solutions cause much difficulty for readers.

We illustrate with a specific example. Suppose one has theoretical grounds to test whether a given distribution has a new mean greater than 10 when the previous mean was 10. These hypotheses can be written either as

$$H_0:\ \mu = 10 \quad H_A:\ \mu > 10 \qquad (1)$$

or

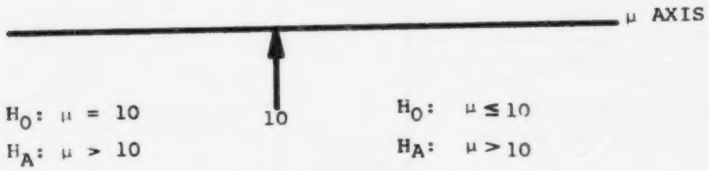
$$H_0:\ \mu \leq 10 \quad H_A:\ \mu > 10 \qquad (2)$$

where $\leq$ is read less than or equal.

The parameter space for this one-sided test is shown in Figure 1. In both (1) and (2) the alternative hypothesis occupies one side of the parameter space line, and hence its name. On the otherhand the null hypothesis is only a point for (1) but a half line for (2).

Figure 1

PARAMETRIC SPACE FOR A ONE-SIDED TEST



---

*Mathematical Statistician, Special Training Branch, Training Institute, ECA.

It is logical to define the hypotheses as (1) since the definition of α remains as the probability of rejecting H_0 if true. It is very unsatisfactory when one realizes that if the true mean is less than 10 the correct decision is impossible. A satisfying decision procedure would be one which makes a correct decision for all possible values of the parameter. This is true only when the union of the hypotheses is the entire parameter line.

It is satisfying to define the hypotheses as (2) but illogical since the α error must be redefined as the probability of rejecting H_0 when $\mu = 10$ only. This follows since the null hypothesis in (2) has a infinite number of different values for μ. Hence for a given α there are an infinite number of critical values with a corresponding number of acceptance and rejection regions. Out of this set there is only one which gives H_A its greatest power and that is when $\mu = 10$. This is the one used in a test and hence the redefinition of α.

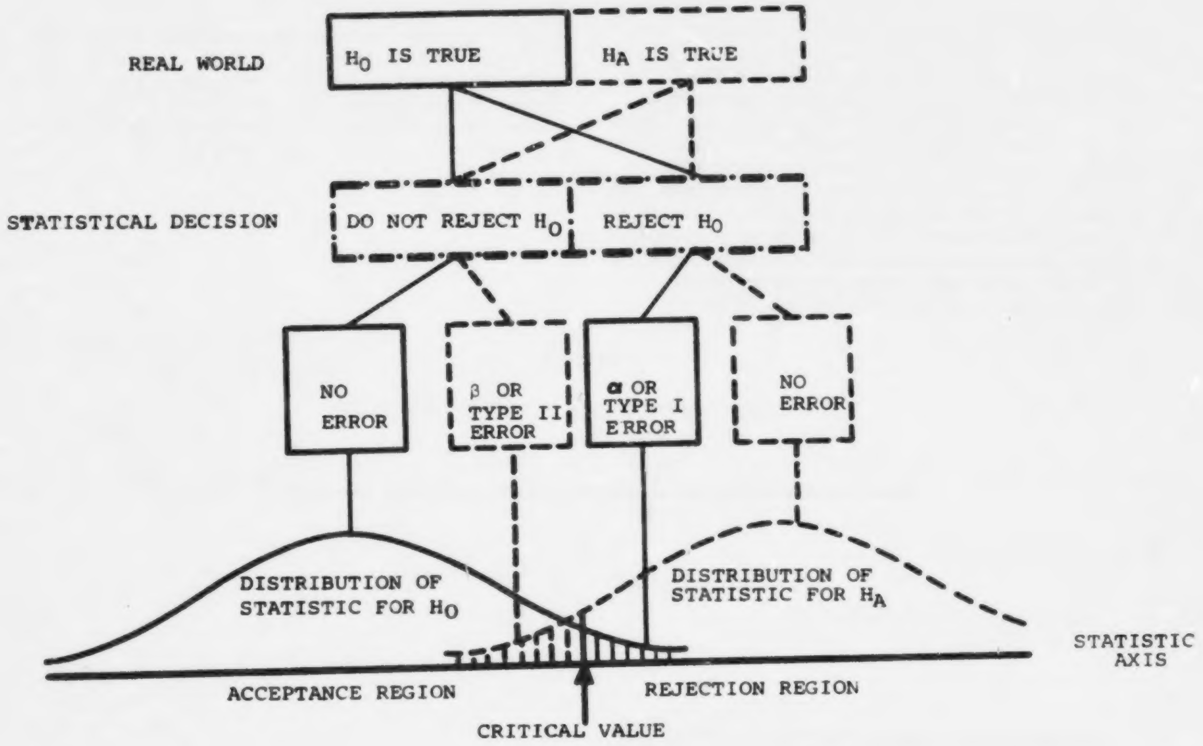
Those who prefer (1) object to the illogical redefinition of α and rightly so. Those who are partial to (2) object to unsatisfactory incomplete specification of the parameter space. The saving grace is that, regardless of which specification is used, the program of action and final decisions agree either that there is not enough evidence to support the research hypothesis or that the research hypothesis has been proved statistically. I prefer (2) and shall use it throughout with α redefined.

III DIAGRAM OF A ONE-SIDED SIGNIFICANCE TEST

Figure 2 shows the diagram of a one-sided test. Note that the distribution for H_0 is a unique curve and would correspond to H_0 in (1). If (2) were used then the distribution of H_0 would be the same and obtained by using the $\mu = 10$ value of H_0.

Figure 2

ONE-SIDED SIGNIFICANCE TEST


REAL WORLD
H0 IS TRUE
HA IS TRUE
STATISTICAL DECISION
DO NOT REJECT H0
REJECT H0
NO ERROR
β OR TYPE II ERROR
α OR TYPE I ERROR
NO ERROR
DISTRIBUTION OF STATISTIC FOR H0
DISTRIBUTION OF STATISTIC FOR HA
STATISTIC AXIS
ACCEPTANCE REGION
REJECTION REGION
CRITICAL VALUE

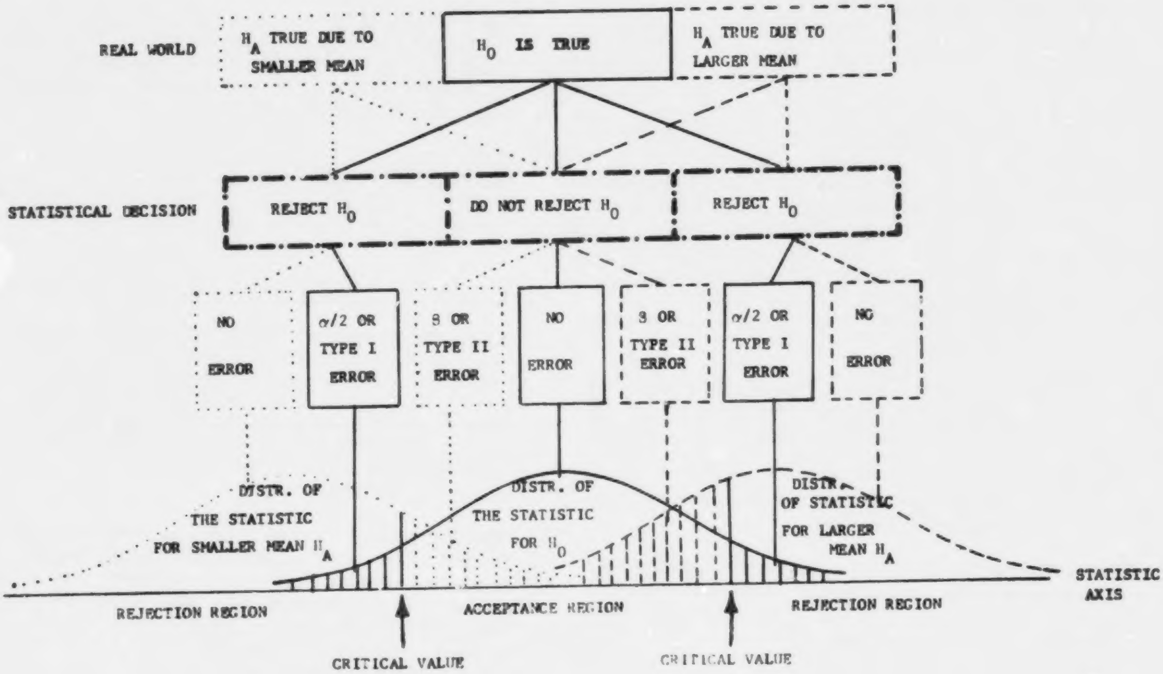
In comparing Figure 2 with the diagram of a two-sided test of Figure 3 some differences are obviously apparent. The most notable is that all of α is in the right tail of H_0 in Figure 2 as opposed to $\alpha/2$ in Figure 3.

This shifting is heuristically obtained by modifying Figure 3 sufficiently to obtain a one-sided test. Since larger differences are the research hypothesis and smaller differences are of no interest the latter distribution should be deleted from Figure 3.

The probability of accepting H_A when true is the area under its curve over the rejection region, and hence the rejection region should be selected to maximize this area. This is done by inspection of Figure 3. It is easily seen that increasing the right rejection region increases power. This is done by moving the right critical value as far to the left as possible (such that all of α is in the right tail of H_0), which maximizes power.

Figure 3

TWO-SIDED TEST OF SIGNIFICANCE


REAL WORLD
H_A TRUE DUE TO SMALLER MEAN
H_0 IS TRUE
H_A TRUE DUE TO LARGER MEAN
STATISTICAL DECISION
REJECT H_0
DO NOT REJECT H_0
REJECT H_0
NO ERROR
$\alpha/2$ OR TYPE I ERROR
β OR TYPE II ERROR
NO ERROR
β OR TYPE II ERROR
$\alpha/2$ OR TYPE I ERROR
NO ERROR
DISTR. OF THE STATISTIC FOR SMALLER MEAN H_A
DISTR. OF THE STATISTIC FOR H_0
DISTR. OF STATISTIC FOR LARGER MEAN H_A
STATISTIC AXIS
REJECTION REGION
ACCEPTANCE REGION
REJECTION REGION
CRITICAL VALUE
CRITICAL VALUE

Note that this shifting or increasing the size of the right rejection region in turn increases the area under H_A over the rejection region or increases the probability of accepting H_A when true or increases the power of the test. Therefore if H_A is true the one-sided test has more power than a corresponding two-sided test, or in other words, one-sided test are preferred.

The discussion of Figure 2 follows analogously that of Figure 3 in part one of this paper and is left to the reader.

IV EXAMPLE OF A ONE-TAILED t-TEST

An example of a one-sided test could be as follows. There exists theoretical grounds for testing whether a new product average exceeds the previous standard = 7.35. The observational data are random and independently normally distributed X_i $i = 1, 2, \ldots, 25$.

Step 1. H_0: $\mu \leq 7.35$

H_A: $\mu > 7.35$

$\alpha = 0.05$

Step 2. The statistic $= \sqrt{n}(\overline{X} - 7.35)/s$, which distributed as t-distribution with degrees of freedom = 24. From a set of t tables, the critical value is 1.711 (see Figure 2). The acceptance region is less than or equal to 1.711, and the rejection region is all other values.

Step 3. From the data $\overline{X} = 7.6$ and $s = 0.504$; hence the statistic $= \sqrt{25}(7.60 - 7.35)/0.502 = 2.48$.

Step 4. Reject H_0 since the statistic is in the rejection region, or the new average exceeds the old.

Modifying the test for a research hypothesis with a smaller mean would be as follows:

Step 1. H_0: $\mu \geq 7.35$

H_A: $\mu < 7.35$

$\alpha = 0.05$

Step 2. The statistic is the same as in the previous problem. Critical value $= -1.711$ with the acceptance region greater than or equal to -1.711 and the rejection region all other values.

Step 3. From the data $\overline{X} = 7.10$, $s = 0.504$; hence the statistic $= \sqrt{25}(7.10 - 7.35)/0.504 = -2.48$.

Step 4. Reject H_0

V SIMPLE AND COMPOSITE HYPOTHESES

At this time it might be well to define two other terms used in connection with hypothesis testing. A simple hypothesis is one that specifies the distribution uniquely or is a point in the parameter space of the distribution. A composite hypothesis is one that is not simple.

If the distribution for Figure 1 has one parameter, say the mean, then H_0 as written on the left is a simple hypothesis or a point. H_0 on the right is composite or the half line while H_A is composite in both statements.

An example of a two-parameter family would be the normal distribution as shown in Figure 4. If the test H_0: $\mu = 10$ is made with known variance = 10 then H_0 is a simple hypothesis since it is the point (10, 10).

If the test is H_0: $\mu = 5$ with an unknown variance then H_0 is composite or the half line $(5, \sigma^2)$ or the unmarked, solid, vertical line of Figure 4. The alternative in both cases is composite.

VI VARIATIONS ON A TWO-SIDED t-TEST

When one reads two different texts giving the same test, confusion can result if the test is presented in what looks like two different programs of action. The confusion is muddied even further when more references are consulted. Figure 5 shows some common ways of presenting a two-sided t-test for the mean where $S_{\overline{x}} = s/\sqrt{n}$.

The top line of Figure 5 agrees with Figure 3. Note that one can algebraically manipulate from one line of Figure 5 to another. In some cases it is possible to do a two-sided test with only one critical value, as shown in line 2 A good example of a two-sided test with only one possible critical value is the F-test used in an analysis of variance. The right tail is always used since we have the ratio of a numerator which consists of a larger or equal mean square divided by a denominator mean square.

Note that lines 3 and 4 have no statistic or critical values as defined in this paper. They would be classified as quasi-test of significance by this writer.

Figure 4

PARAMETER SPACE FOR THE NORMAL DISTRIBUTION

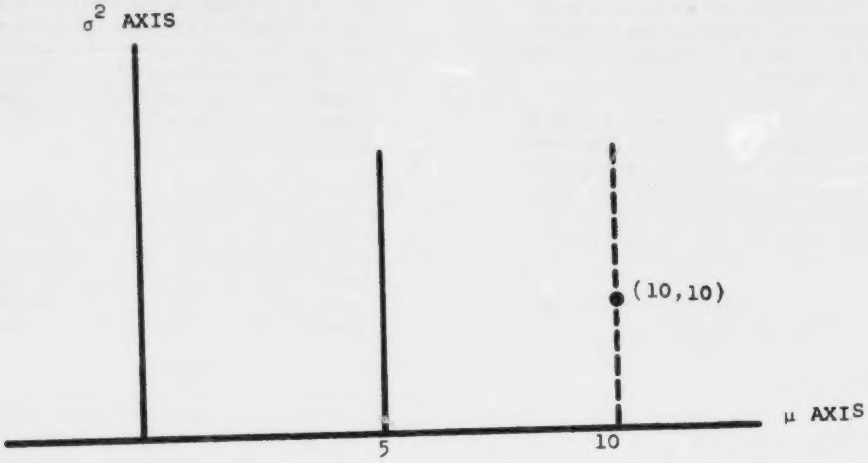


Figure 5

VARIATIONS ON A TWO-SIDED TEST

EXPLANATION	METHOD OF TEST	COMMENTS
RANDOM VARIABLE AND TEST STATISTIC	$-t < (\bar{X} - \mu)/S_{\bar{X}} < t$	CRITICAL VALUES t & -t
RANDOM VARIABLE AND TEST STATISTIC	$\lvert \bar{X} - \mu \rvert / S_{\bar{X}} < t$	CRITICAL VALUE t
RANDOM VARIABLES NO TEST STATISTIC	$-tS_{\bar{X}} < (\bar{X} - \mu) < tS_{\bar{X}}$	QUASI CRITICAL VALUES $\pm tS_{\bar{X}}$
RANDOM VARIABLES NO TEST STATISTIC	$\lvert \bar{X} - \mu \rvert < tS_{\bar{X}}$	QUASI CRITICAL VALUE $tS_{\bar{X}}$
RANDOM VARIABLES NO TEST STATISTIC	$\bar{X} - tS_{\bar{X}} < \mu < \bar{X} + tS_{\bar{X}}$	CONFIDENCE LIMITS FOR μ

At times it may be possible to manipulate a significance test into a confidence limits. Hence one can perform a test by using confidence limits, as shown in the last line of Figure 5. This result is not always true. Here estimation and tests are on common ground. Lines 1 and 5 of Figure 5 are mathematically equivalent.

If the reader is unsure of Figure 5 then a numerical example should be used to work out all cases. If the previous numerical example is used then arithmetic and time are at a minimum with the advantage that the statistical decision can be checked for agreement if $\alpha = 0.1$ is used.

Variations on a one-sided test are left to the reader. They can be easily obtained from lines 1, 3, and 5 with the appropriate modification.

VII CONCLUDING REMARKS

We conclude with several remarks that should help clarify further the ideas associated with a test of significance.

A For small samples a significant result will be obtained only if the null hypothesis is very badly violated. In Figure 2 this means that the distance between the peaks of H_0 and H_A must be very large. Small samples imply large variances for both distributions, and hence poor power. It is a truism that from small samples you get small information.

B If a difference exists and is shown, then by increasing sample size this difference can be shown with a smaller type-I error. For example, if one gets a significant result at $\alpha = 0.05$ and a sample size of 20, then with a sample size of 50 a significant result can be obtained with a smaller α. The effect of increasing sample size in Figure 2 would be to cause the variance of both H_0 and H_A to be decreased or both distributions would be more peaked. Hence to keep α fixed the critical value must be shifted to the left, and this shift thereby increases the rejection region or power of the test. In addition since H_A is peaked the type-II error would be decreased without changing the critical value, and hence power increases for a second reason. The sum of these two increments gives a more powerful test. Therefore α can be decreased or the critical value moved right to retain the original power, which was significant.

C Using the same α and increasing sample size will result in a rejection of H_0 even when differences are smaller and thus more difficult to detect. As an example, if $\alpha = 0.05$ and sample size = n failed to get a significant result because of a type-II error, then increasing the sample size sufficiently will result in a significant statistic. Increasing the sample size increases the power for the same α, as discussed in B above.

D For the same sample size and α, larger mean differences are more easily discriminated. As the difference between H_0 and H_A increases, this means in Figure 2 that the peaks of the two distributions become farther apart. This increases power and the probability of a significant result.

E For the same sample size and α, smaller differences are more difficult to determine. The heuristic argument follows that in the pervious paragraph, with H_A and H_0 shifted closer together, loss in power can be regained by increasing sample size.

F It is often claimed that if a difference exists, no matter how small, one can prove it provided that sample size is large enough. One can reduce both type-I and type-II errors by increasing sample size. If sample size is infinite then both errors can be made zero.

G If after increasing sample size the same decision of nonsignificance is obtained the confidence of agreement between the sampled and hypothetical population is strengthened. In other words, if one continuously gets nonsignificant results with increasing sample size, then the acceptance of H_0 becomes more tenable. Complete confidence in H_0 is obtained only with an infinitely large sample.

H Trying to prove the validity of H_0 by nonsignificant results for finite n is like the mathematician's trying to prove a theorem by example. If the mathematician's theorem is true for every possible example, then it is proved. By the same reasoning, if every possible sample of size n is tested then one would expect to have a significant result of αN times, where N = total number of all possible samples of size n. Significant deviations from αN would invalidate H_0.

I If a test does not meet the assumptions then the α level is affected and is not what the researcher specified. Power may also be adversely affected. A test with fewer assumptions is preferred unless the departures are compensated for by robustness.

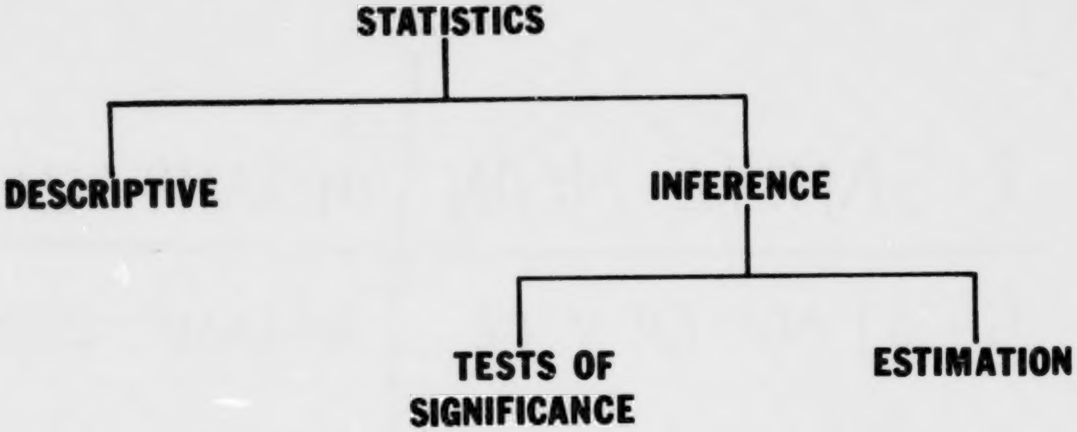
J If a researcher really understood the test he was using, with enough time he could determine the critical value empirically via a Monte Carlo simulation study. Repeated sampling would permit the construction of the distribution of the statistic for H_0. From it the critical value could be easily determined. If the researcher cannot imagine an experiment for obtaining the distribution via some random sampling procedure, then he really does not understand his test of significance.

K Similarly, by using the Monte Carlo technique, one can empirically determine the power of the test versus that of a specific H_A. First assume that H_A has a given set of parameter values, next determine its approximate distribution via Monte Carlo, and determine $1-\beta$ or the power of the area under the curve over the rejection region. By assuming a different set of parameter values and by repeating the above, one can study the change in the power of a test versus any set of specified parameter values.

L In the same way one can use Monte Carlo techniques to investigate robustness. If one assumes a departure of a certain magnitude from the basic assumption, then the evaluation of α is determined as in J, where the rejection region is now known from the distribution of H_0 when no assumptions are violated. If these two α's are in agreement, then the procedure is robust with respect to the proceeding departures from the basic assumptions. This procedure can be repeated for any assumption or combination of them or any magnitude of departure from them.

STATISTICAL DATA

Classification of Statistic


STATISTICS
DESCRIPTIVE
INFERENCE
TESTS OF SIGNIFICANCE
ESTIMATION

Tests of significance (hypothesis testing, tests of hypothesis, significance tests, ...)

A test of significance can be defined as a method of analyzing data so as to discriminate between two hypothesis.

1) PROBLEM DEFINITION

2) DATA COLLECTION SCHEME

3) " "

4) " ANALYSIS

5) WRITTEN REPORT

A statistic (note no s on the end) can be simply defined as a function of the sample and is the measure or criterion used to discriminate between two hypotheses used in a test of significance, namely the null and alternative hypothesis.

$\bar{X}$ = SAMPLE MEAN

n = SAMP. SIZE

5 = STAN. DEV. OF PAR. POP.

s = SAMP. EST. OF STAN. DEV.

1) $z = \sqrt{n}\,(\bar{x} - 10)/5$ 2) $t = \sqrt{n}\,(\bar{x} - 10)/s$

EX. 10, 11, 12, 11 $\bar{x} = 44/4 = 11$

$z = \sqrt{4}\,(11-10)/5 = 2/5 = 0.4$

Two examples of a statistic are formula 1) and 2) above.

The null hypothesis is defined as the hypothesis of no difference between a hypothetical and the sampled population.

$H_0: P_S = P_H$ OR $H_0: P_S - P_H = 0$

1) $Z = \sqrt{n}\,(\bar{X}-10)/5$

$H_0: \mu = 10$

Hypothetical population (P_H) is normally distributed with mean = 10 and standard deviation = 5.

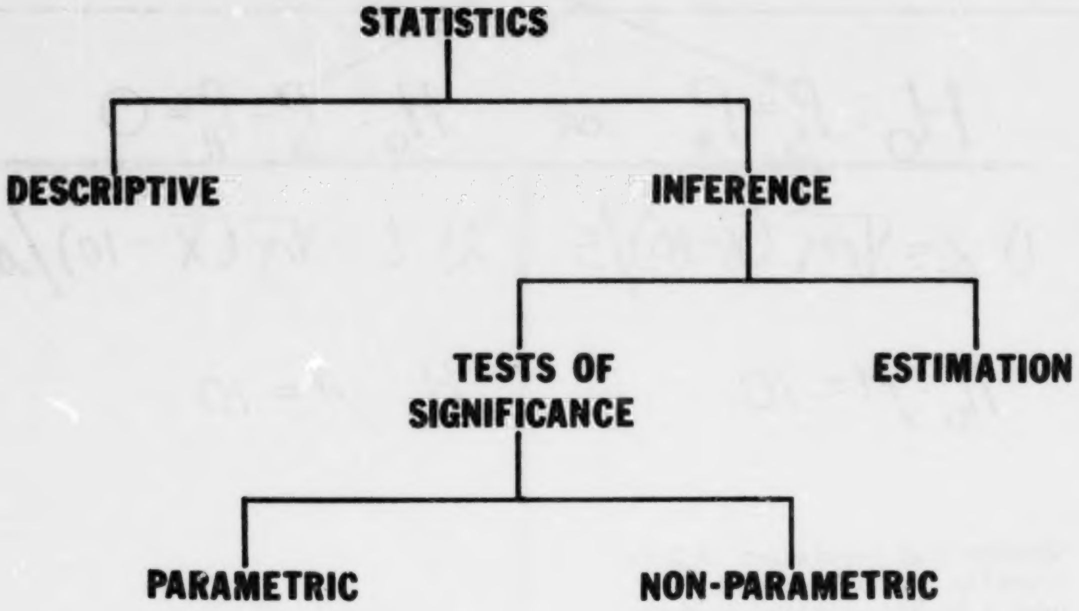
Sampled population (P_S) is normally distributed with unknown mean and standard deviation = 5

2) $t = \sqrt{n}\,(\bar{X}-10)/s$

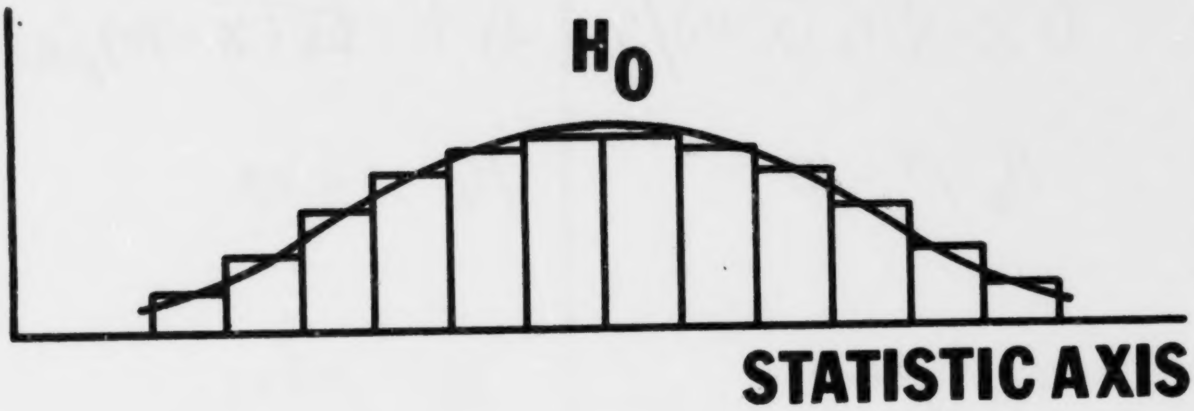
$H_0: \mu = 10$

Each test of significance has associated with it explicit or implicit, or both, test assumptions that must be satisfied if the results of the test are to be valid.

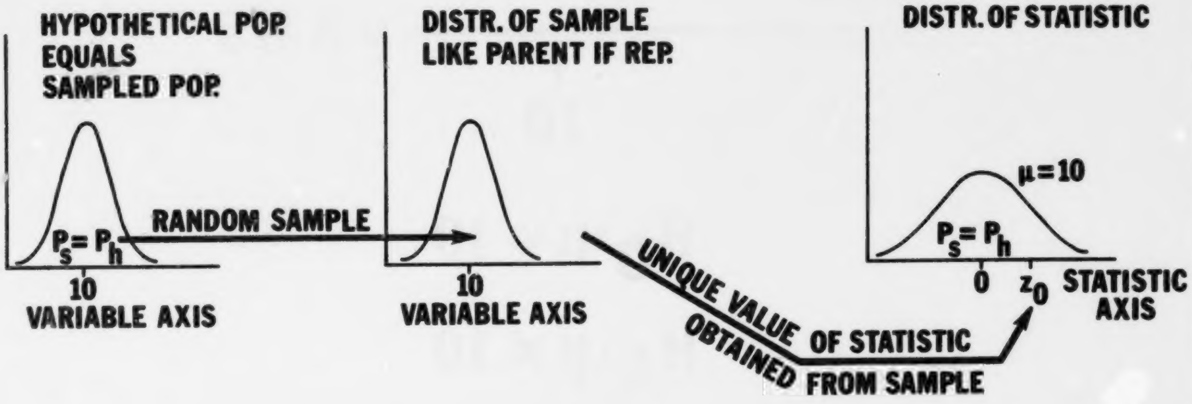
Classification of Statistic


STATISTICS
DESCRIPTIVE
INFERENCE
TESTS OF SIGNIFICANCE
ESTIMATION
PARAMETRIC
NON-PARAMETRIC

Distribution of the Statistic When the Null Hypothesis Is True



Four Distributions in Hypothesis Testing
H_O is True

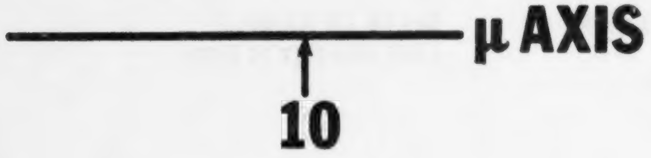

HYPOTHETICAL POP. EQUALS SAMPLED POP.
$P_s = P_h$
10
VARIABLE AXIS
RANDOM SAMPLE
DISTR. OF SAMPLE LIKE PARENT IF REP.
10
VARIABLE AXIS
UNIQUE VALUE OF STATISTIC OBTAINED FROM SAMPLE
DISTR. OF STATISTIC
$\mu=10$
$P_s = P_h$
0
z_0
STATISTIC AXIS

The alternative hypothesis is usually the negation of the null hypothesis but can be any admissible hypothesis alternative to the one under the test.

$$H_0 : P_S = P_H \quad \text{OR} \quad H_0 : P_S - P_H = 0$$

1) $Z = \sqrt{n}\,(\bar{X} - 10)/5$	2) $t = \sqrt{n}\,(\bar{X} - 10)/s$
$H_0 : \mu = 10$	$H_0 : \mu = 10$
$H_A : \mu \neq 10$	$H_A : \mu \neq 10$

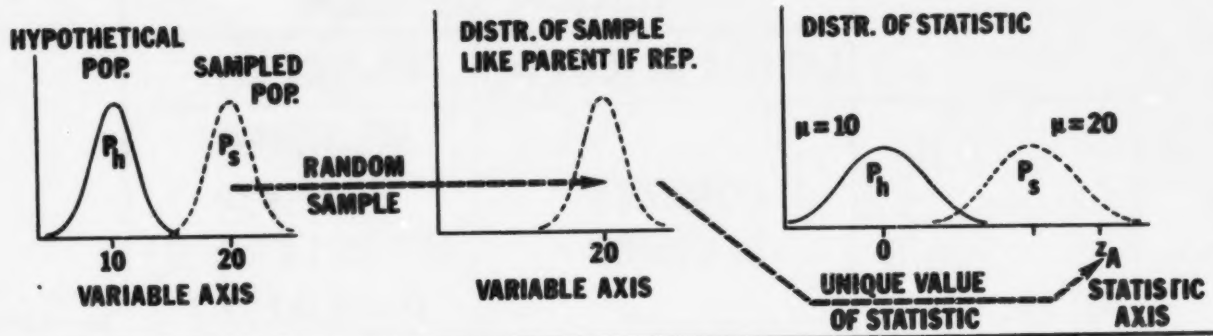
Parametric Space for a Two-Sided Test



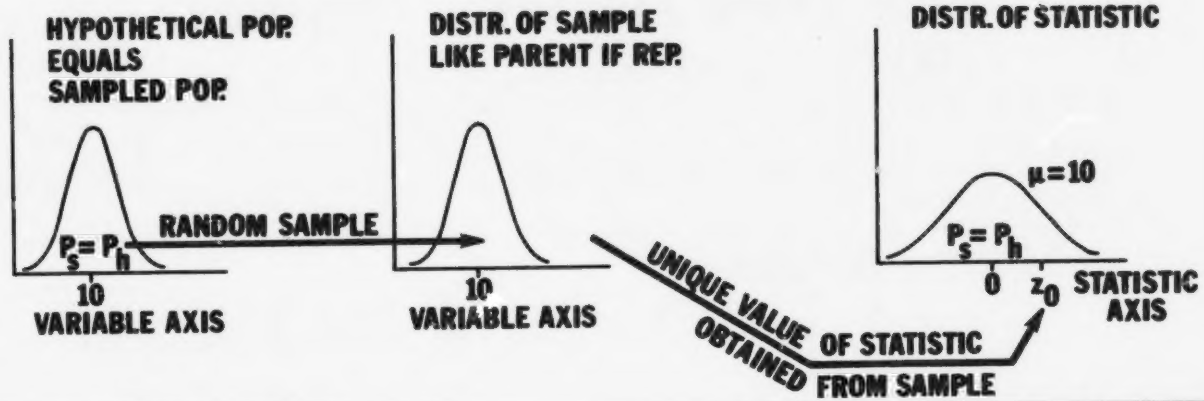
$H_0 : \mu = 10$

$H_A : \mu \neq 10$

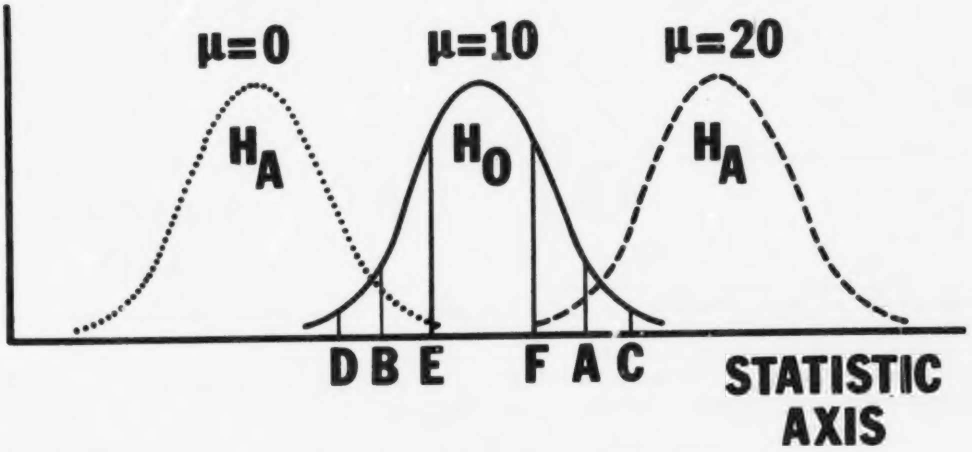
Four Distributions in Hypothesis Testing
H_A is True


HYPOTHETICAL POP.
SAMPLED POP.
P_h
P_s
RANDOM SAMPLE
10
20
VARIABLE AXIS
DISTR. OF SAMPLE LIKE PARENT IF REP.
20
VARIABLE AXIS
DISTR. OF STATISTIC
$\mu=10$
$\mu=20$
P_h
P_s
0
z_A
UNIQUE VALUE OF STATISTIC
STATISTIC AXIS

Four Distributions in Hypothesis Testing
H_0 is True


HYPOTHETICAL POP. EQUALS SAMPLED POP.
$P_s = P_h$
RANDOM SAMPLE
10
VARIABLE AXIS
DISTR. OF SAMPLE LIKE PARENT IF REP.
10
VARIABLE AXIS
DISTR. OF STATISTIC
$\mu=10$
$P_s = P_h$
0
z_0
STATISTIC AXIS
UNIQUE VALUE OF STATISTIC OBTAINED FROM SAMPLE

Some Possible Distribution for the Statistic


$\mu=0$
$\mu=10$
$\mu=20$
H_A
H_0
H_A
D
B
E
F
A
C
STATISTIC AXIS

Four Possible Decisions in a Test of Significance

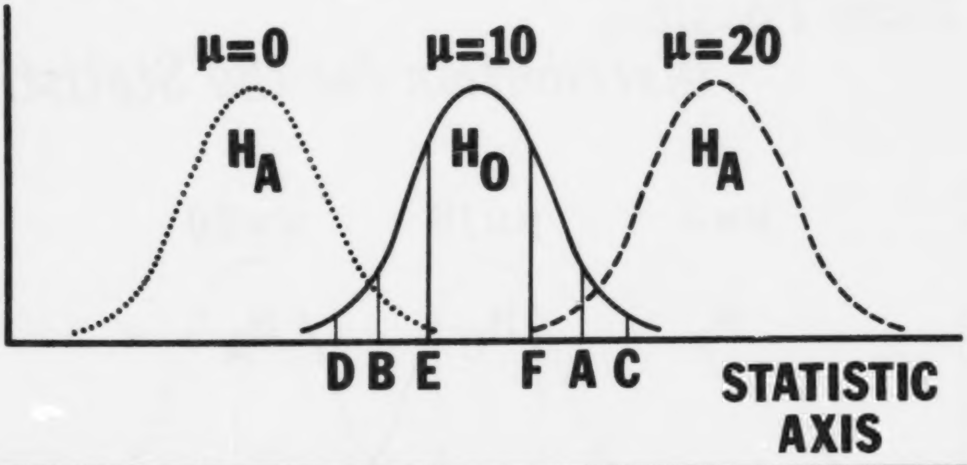
		STATISTICAL DECISION	
		DO NOT REJECT H_O	REJECT H_O
REAL WORLD	H_O IS TRUE	CORRECT DECISION	INCORRECT DECISION (TYPE I ERROR) (α-ERROR)
	H_O IS NOT TRUE	INCORRECT DECISION (TYPE II ERROR) (β-ERROR)	CORRECT DECISION

Error of the First Kind

Error of the Second Kind

α-error (Type I Error, Error of First Kind) is the rejection of the null hypothesis when it is true.

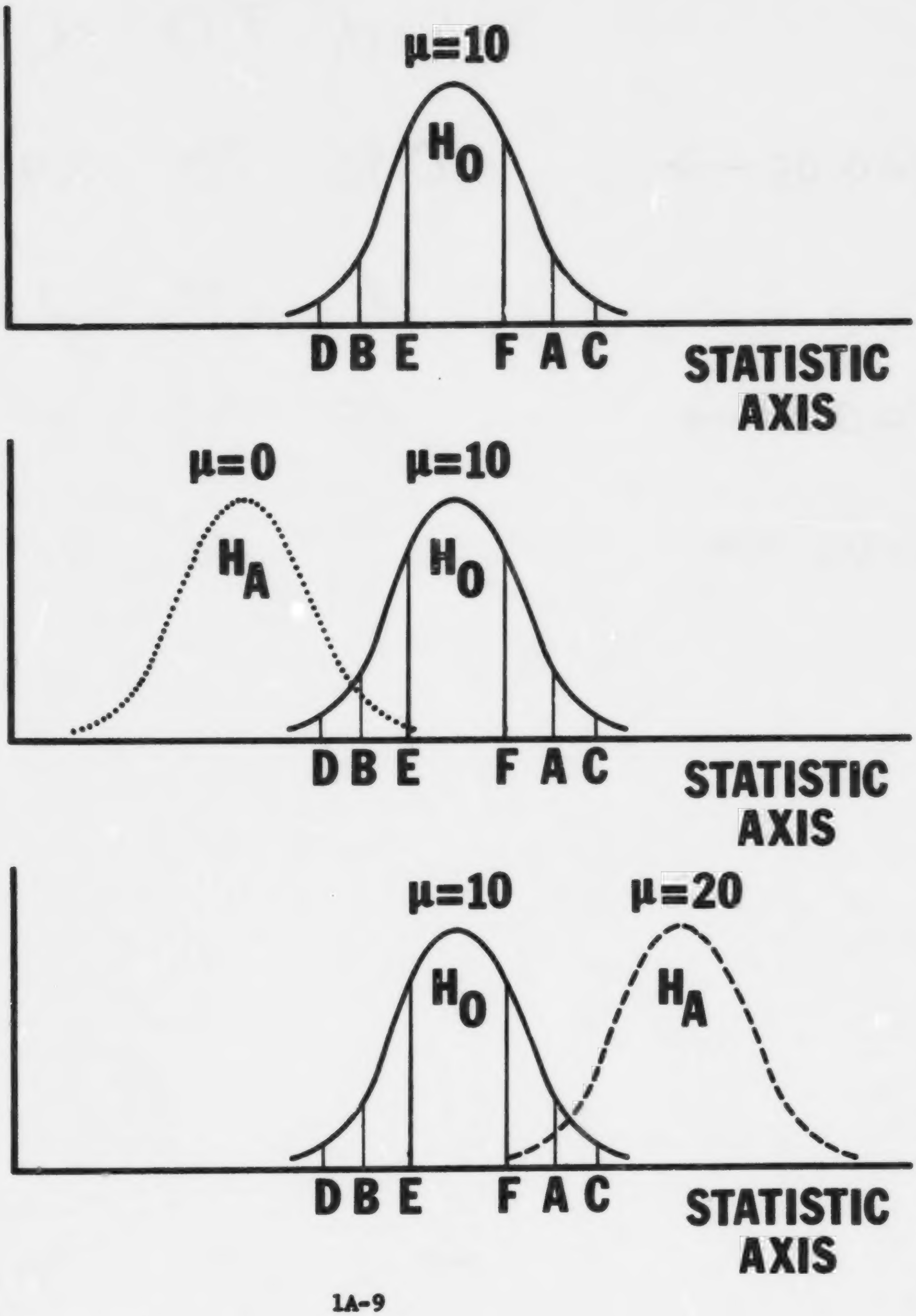
Some Possible Distribution for the Statistic



The magnitude of the α-error is called α and is also known as the level of significance.

The critical values (or value) partition the statistics axis into two regions, the acceptance region and the rejection region.

Some Possible Distribution for the Statistic


μ=10
H0
D B E F A C
STATISTIC AXIS
μ=0
μ=10
HA
H0
D B E F A C
STATISTIC AXIS
μ=10
μ=20
H0
HA
D B E F A C
STATISTIC AXIS

ODDS OF CORRECT DEC.-H_0 TRUE

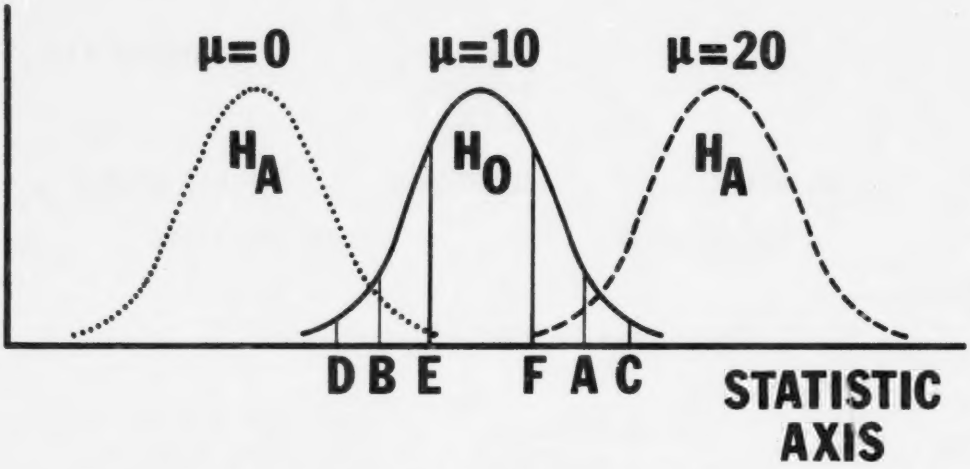
	$1-\alpha$ TO α
$\alpha = 0.05$ →	0.95 TO 0.05
	19 TO 1
$\alpha = 0.01$ →	99 TO 1
$\alpha = 0.1$ →	9 TO 1

β-error (Type II Error, Error of Second Kind) is not rejecting the null hypothesis when it is false, or is the rejection of the alternative hypothesis when it is true

The magnitude of the β-error is called β.

The α-error is generally considered more serious than a β-error.

Some Possible Distribution for the Statistic



One minus beta (1 - β) is defined as the power of the test.

The odds of accepting the alternative hypothesis (H_A) when true are 1 - β to β.

Parametric test are generally more powerful than the corresponding non-parametric test.

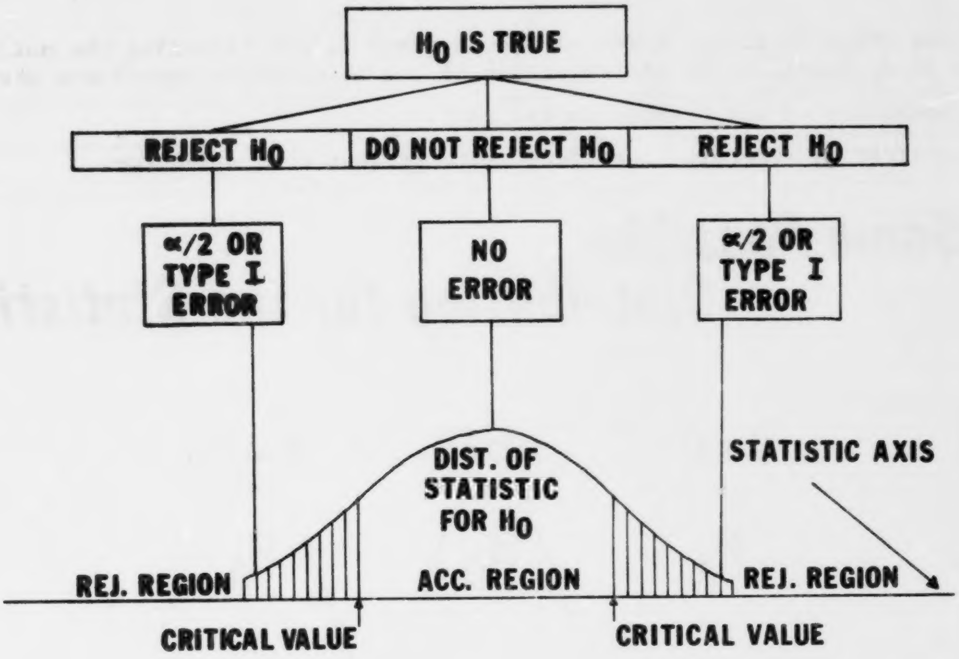
The rejection region is a region on the statistic axis such that if the statistic fall within it the null hypothesis is rejected. It is also called significance region and critical region.

The acceptance region is a region on the statistic axis such that if the statistic falls within it the null hypothesis is not rejected.

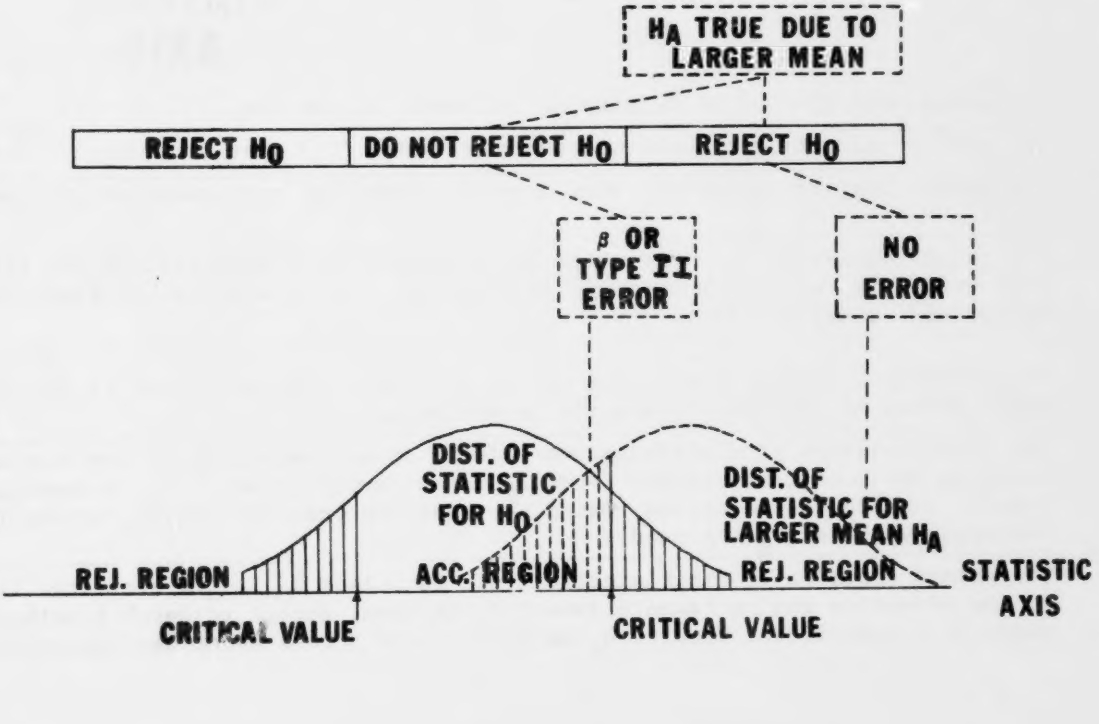
The first decision or discrimination rule is do not reject H_O if the statistic falls in the acceptance region (some aliases: cannot reject H_O, a non-significant result, outside the rejection region, outside the critical region, accept H_O, research hypothesis not proved).

The second decision or discrimination rule is reject H_O if the statistic falls in the rejection region (some aliases: H_A is true, accept research hypothesis, there is a significant result, H_A is true unless a rare event has happened).

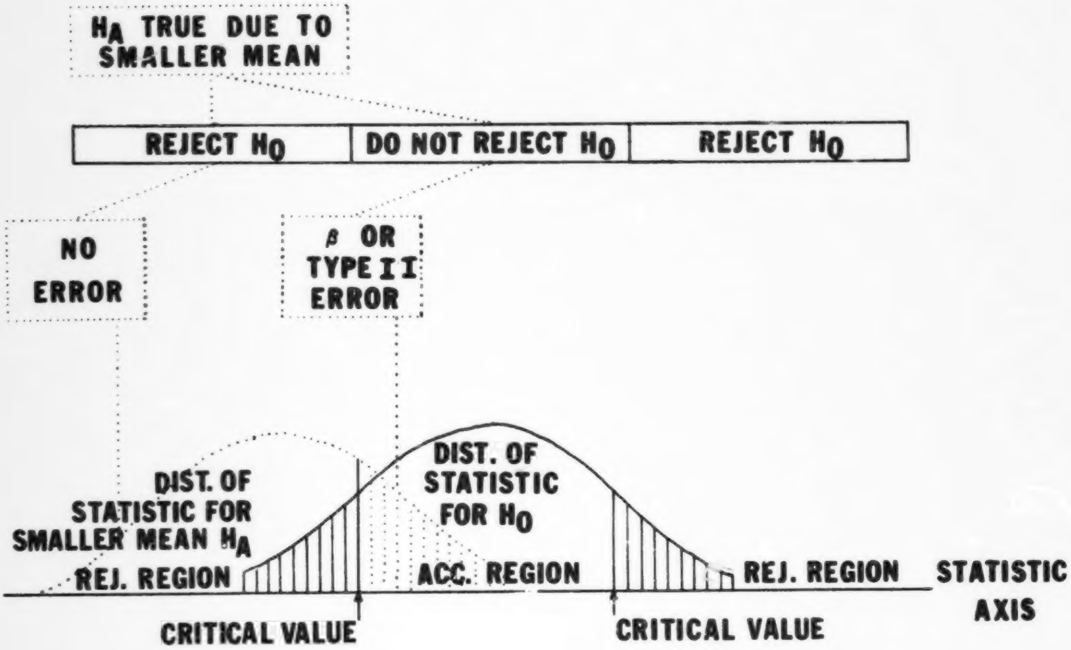
Two·Sided Test of Significance


H_0 IS TRUE
REJECT H_0
DO NOT REJECT H_0
REJECT H_0
$\alpha/2$ OR TYPE I ERROR
NO ERROR
$\alpha/2$ OR TYPE I ERROR
DIST. OF STATISTIC FOR H_0
STATISTIC AXIS
REJ. REGION
ACC. REGION
REJ. REGION
CRITICAL VALUE
CRITICAL VALUE

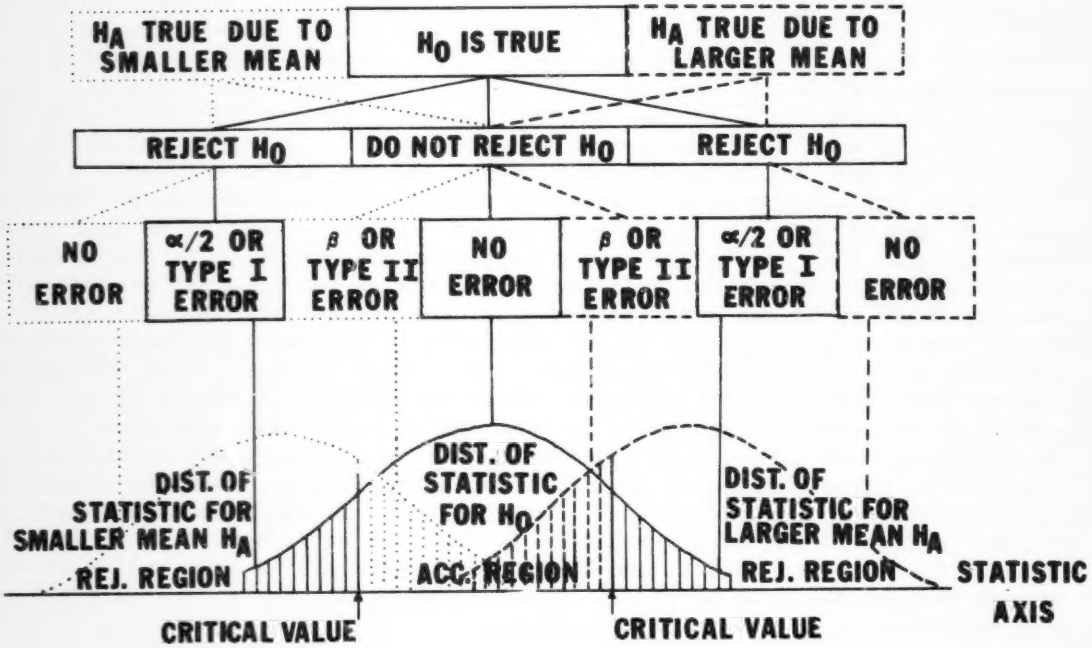
Two·Sided Test of Significance


H_A TRUE DUE TO LARGER MEAN
REJECT H_0
DO NOT REJECT H_0
REJECT H_0
β OR TYPE II ERROR
NO ERROR
DIST. OF STATISTIC FOR H_0
DIST. OF STATISTIC FOR LARGER MEAN H_A
REJ. REGION
ACC. REGION
REJ. REGION
STATISTIC AXIS
CRITICAL VALUE
CRITICAL VALUE

Two-Sided Test of Significance


H_A TRUE DUE TO SMALLER MEAN
REJECT H_0
DO NOT REJECT H_0
REJECT H_0
NO ERROR
β OR TYPE II ERROR
DIST. OF STATISTIC FOR SMALLER MEAN H_A
DIST. OF STATISTIC FOR H_0
REJ. REGION
ACC. REGION
REJ. REGION
STATISTIC AXIS
CRITICAL VALUE
CRITICAL VALUE

Two-Sided Test of Significance


H_A TRUE DUE TO SMALLER MEAN
H_0 IS TRUE
H_A TRUE DUE TO LARGER MEAN
REJECT H_0
DO NOT REJECT H_0
REJECT H_0
NO ERROR
α/2 OR TYPE I ERROR
β OR TYPE II ERROR
NO ERROR
β OR TYPE II ERROR
α/2 OR TYPE I ERROR
NO ERROR
DIST. OF STATISTIC FOR SMALLER MEAN H_A
DIST. OF STATISTIC FOR H_0
DIST. OF STATISTIC FOR LARGER MEAN H_A
REJ. REGION
ACC. REGION
REJ. REGION
STATISTIC AXIS
CRITICAL VALUE
CRITICAL VALUE

A robust test is one that is insensitive to departures from the assumptions. A test that is sensitive to departures from the test assumptions lacks robustness.

A test of significance can be defined as a method of analyzing data so as to discriminate between two hypotheses. The first hypothesis is the null hypothesis; the second is the alternative hypothesis, which should be the operational statement of the experimenter's research hypothesis. Next the best test for the alternative hypothesis is selected, that is, the one with the greatest power, and is robust. This test is based upon assumptions, which may be explicitly or implicitly stated, or both. The probability statements used in discrimination are based upon the assumptions. After the test is selected, the α-error is prespecified, that is the odds of being right if the null hypothesis is true. This α-error in conjunction with the known distribution of the statistic for the null hypothesis determines critical values. The critical values partition the entire statistic axis into two parts, the acceptance and rejection region. The sample size is fixed and the data are now collected. From the collected data the statistic is evaluated and compared with the critical values to determine whether it is in the acceptance or rejection region. After this comparisem the appropriate decision is made. If the statistic is in the acceptance region, the appropriate decision is do not reject the null hypothesis. If the statistic is in the rejection region, the decision is reject the null hypothesis. With each decision there is a possible error. When the statistic falls in the acceptance region and if in the real world the alternative hypothesis is true and we reject it, this is a Type-II error. On the other hand if the null hypothesis is true and we do not reject it then there is no error. There is no chance for a Type-I error when the statistic is in the acceptance region. When that statistic falls in the rejection region and if in the real world the null hypothesis is true and we reject it, this is a Type-I error. On the other hand if the alternative hypotheis is true and we do not reject it then there is no error. There is no chance for a Type-II error when the statistic is in the rejection region.

X_i ¦ $i = 1, 2, \dots, 25$ ¦ RAN. ¦ IND. ¦ NOR. DISTB.

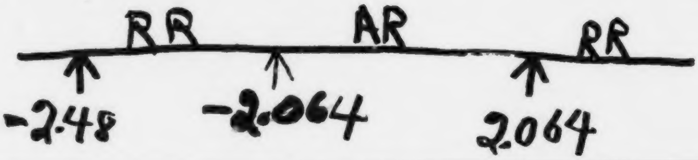
1) $H_0 : \mu = 7.35$ ¦ $H_A : \mu \neq 7.35$ ¦ $\alpha = 0.05$

2) STAT. $= \sqrt{n}\,(\bar{X} - 7.35)/s$ ¦ t-DISTB. ¦ df $= 24$

C.V. $= \pm 2.064$

3) $\bar{X} = 7.1$ ¦ $s = 0.504$ ¦ $t = \sqrt{25}\,(7.1 - 7.35)/.504 = -2.48$

4)



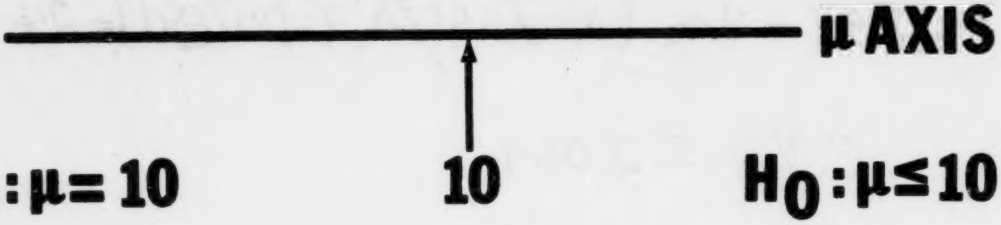
REJ. H_0

If there exists no valid justification (theoretical grounds or previous experience) for a one sided test then a two-sided test is mandatory.

1) $H_0 : \mu = 10$ ¦ $H_A : \mu > 10$

2) $H_0 : \mu \leq 10$ ¦ $H_A : \mu > 10$

Parametric Space for a One-Sided Test



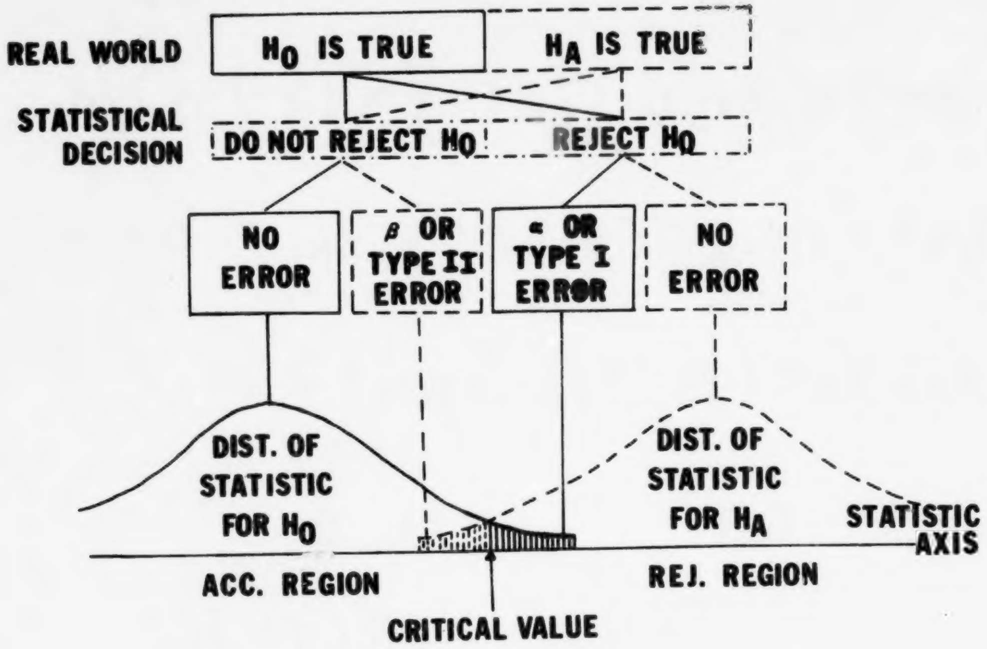
$H_0 : \mu = 10$

$H_A : \mu > 10$

$H_0 : \mu \leq 10$

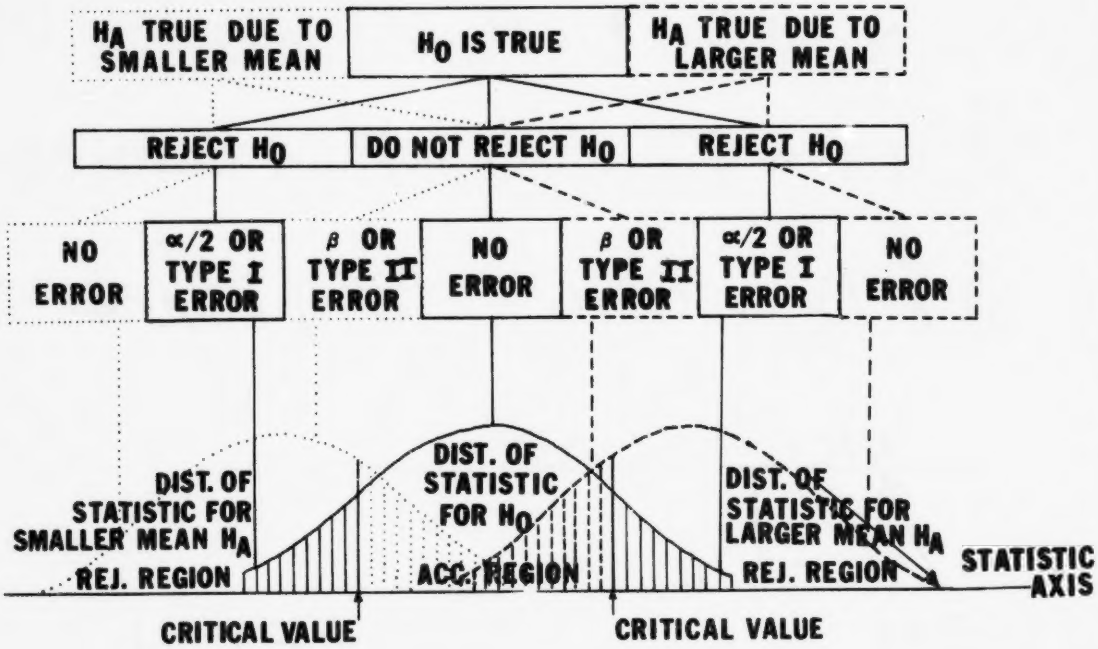
$H_A : \mu > 10$

One-Sided Significance Test



REAL WORLD
H0 IS TRUE
HA IS TRUE
STATISTICAL DECISION
DO NOT REJECT H0
REJECT H0
NO ERROR
β OR TYPE II ERROR
α OR TYPE I ERROR
NO ERROR
DIST. OF STATISTIC FOR H0
DIST. OF STATISTIC FOR HA
STATISTIC AXIS
ACC. REGION
REJ. REGION
CRITICAL VALUE

Two-Sided Test of Significance



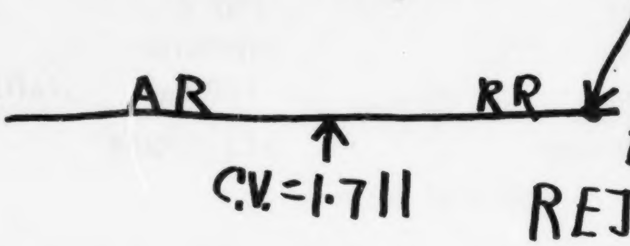
HA TRUE DUE TO SMALLER MEAN
H0 IS TRUE
HA TRUE DUE TO LARGER MEAN
REJECT H0
DO NOT REJECT H0
REJECT H0
NO ERROR
α/2 OR TYPE I ERROR
β OR TYPE II ERROR
NO ERROR
β OR TYPE II ERROR
α/2 OR TYPE I ERROR
NO ERROR
DIST. OF STATISTIC FOR SMALLER MEAN HA
DIST. OF STATISTIC FOR H0
DIST. OF STATISTIC FOR LARGER MEAN HA
REJ. REGION
ACC. REGION
REJ. REGION
STATISTIC AXIS
CRITICAL VALUE
CRITICAL VALUE

X_i ; $i=1,2,\ldots,25$: RAN : IND : NOR : 1) $H_0: \mu \le 7.35$

H_A $\mu > 7.35$ | $\alpha = 0.05$ | STAT $= \sqrt{n}\,(\bar{X}-7.35)/s$

t-DISTB | df=24 | C.V = 1.711 | $\bar{X} = 7.6$ | $s = 0.504$

STAT $= \sqrt{25}\,(7.6-7.35)/0.504 = 2.48$



DEC.

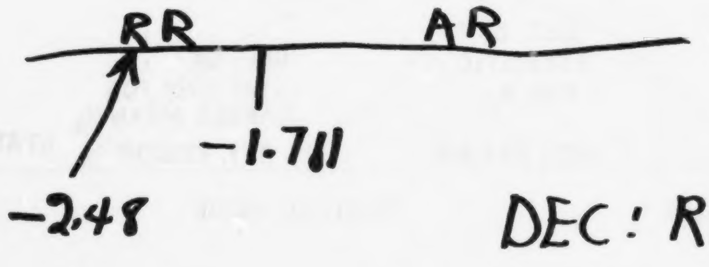
REJ. H_0

---

1) $H_0: \mu \ge 7.35$: $H_A: \mu < 7.35$: $\alpha = 0.05$

2) STAT $= \sqrt{n}\,(\bar{X}-7.35)/s$: t-DISTB : C.V = −1.711

3) $\bar{X} = 7.1$ | $s = 0.504$ | $S = \sqrt{25}\,(7.1-7.35)/0.504 = -2.48$

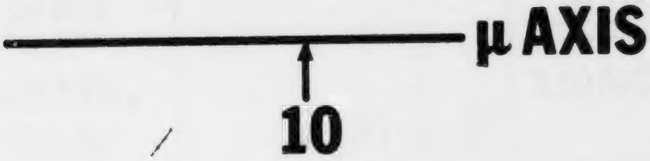


DEC: REJ. H_0

A simple hypothesis is one that specifies the distribution uniquely or is a point in the parameter space of the distribution.

A composite hypothesis is one that is not simple.

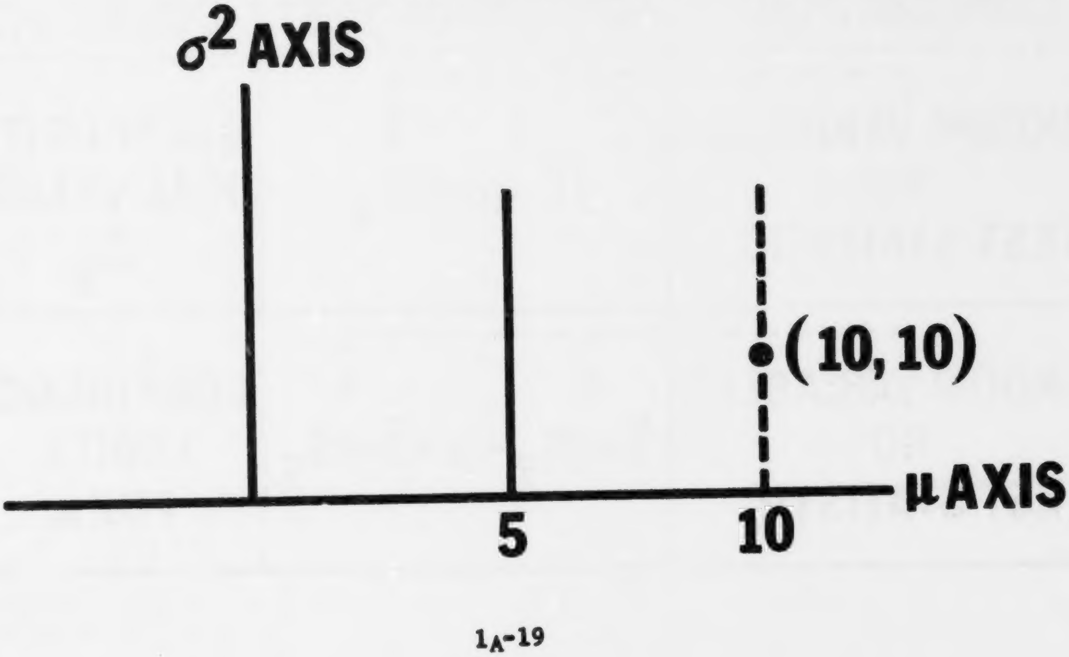
Parametric Space for a Two-Sided Test



$H_0: \mu = 10$

$H_A: \mu \neq 10$

Parameter Space for the Normal Distribution


σ^2 AXIS
(10, 10)
μ AXIS
5
10

Variations on a Two-Sided Test

EXPLANATION	METHOD OF TEST	COMMENTS
RANDOM VARIABLE AND TEST STATISTIC	$-t < (\bar{X}-\mu)/S_{\bar{X}} < t$	CRITICAL VALUE t & -t
RANDOM VARIABLE AND TEST STATISTIC	$\lvert\bar{X}-\mu\rvert/S_{\bar{X}} < t$	CRITICAL VALUE t
RANDOM VARIABLES NO TEST STATISTIC	$-tS_{\bar{X}} < (\bar{X}-\mu) < tS_{\bar{X}}$	QUASI CRITICAL VALUES $\pm tS_{\bar{X}}$

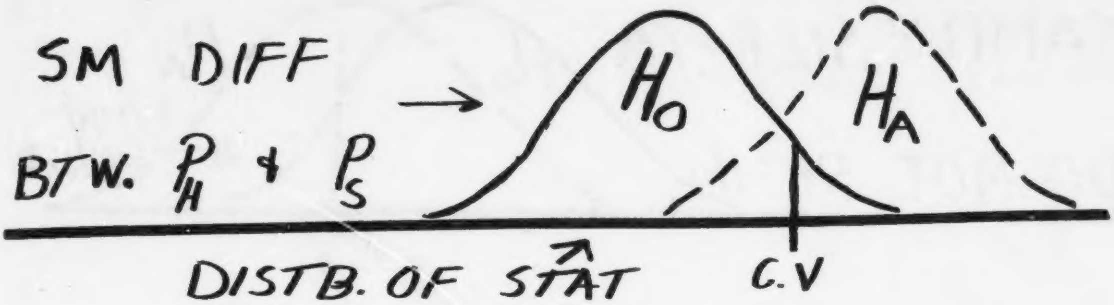
WHERE

$S_{\bar{X}} = \frac{s}{\sqrt{n}}$

Variations on a Two-Sided Test

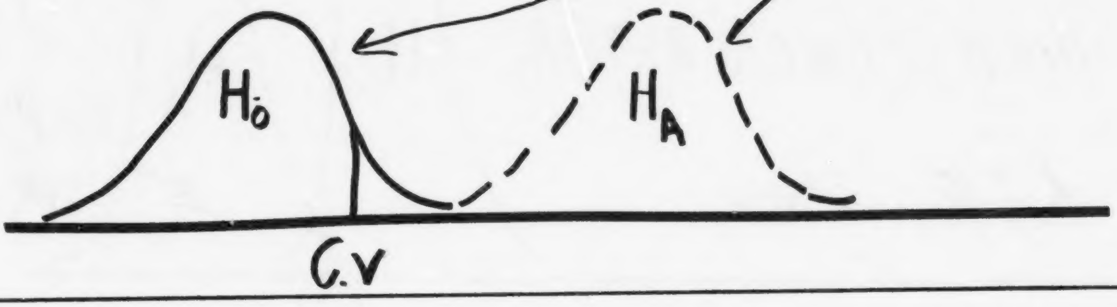
EXPLANATION	METHOD OF TEST	COMMENTS
RANDOM VARIABLES NO TEST STATISTIC	$\lvert\bar{X}-\mu\rvert < tS_{\bar{X}}$	QUASI CRITICAL VALUE $tS_{\bar{X}}$
RANDOM VARIABLES NO TEST STATISTIC	$\bar{X}-tS_{\bar{X}} < \mu < \bar{X}+tS_{\bar{X}}$	CONFIDENCE LIMITS FOR μ

SM DIFF BTW. P_H & P_S →

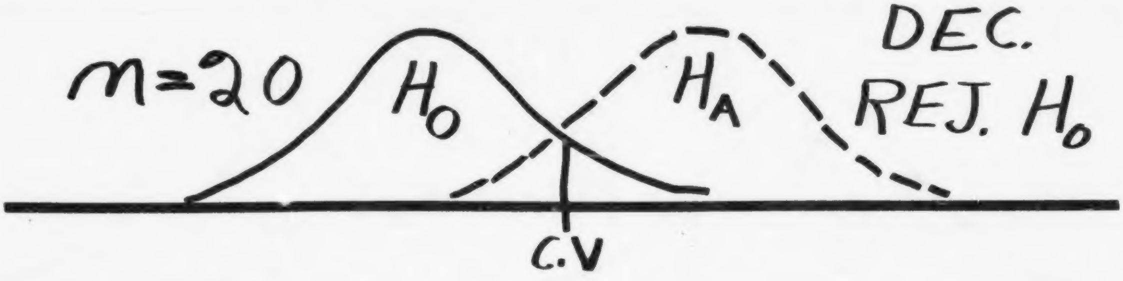


DISTB. OF STAT

LARGE DIFF. BTWN P_H & P_S



$n=20$



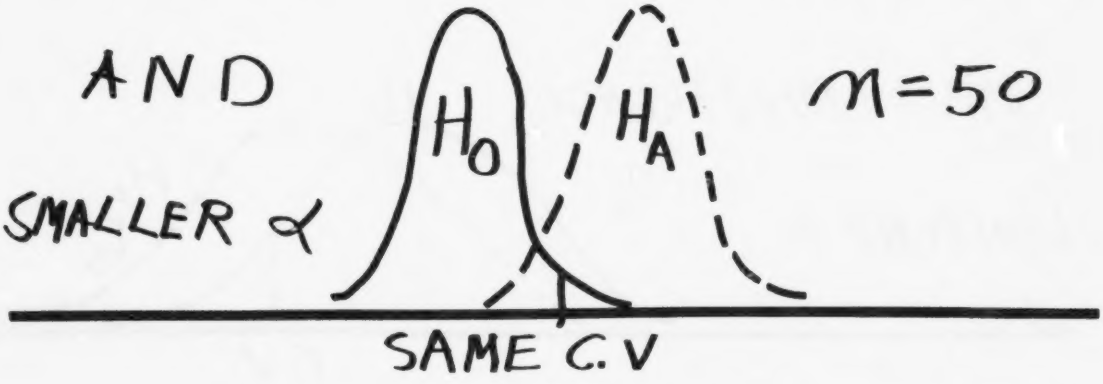
DEC.
REJ. H_0

SAME C.V. → MORE POWER

AND

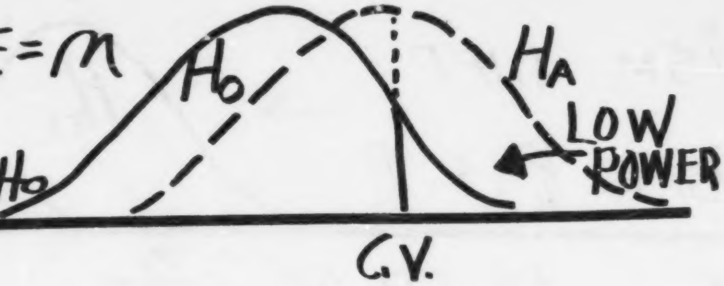
SMALLER α

$n=50$


H_0
H_A
SAME C.V

SAMPLE SIZE=n

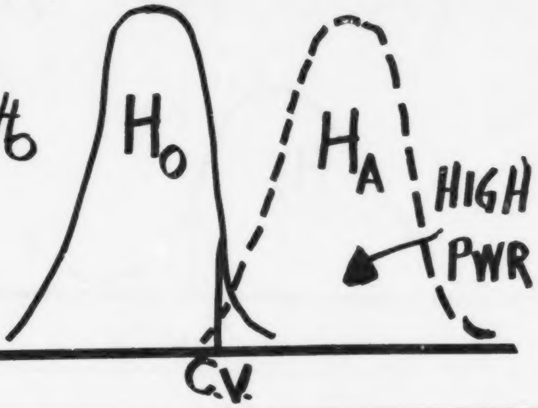
DO NOT REJ H_0


H_0
H_A
LOW POWER
C.V.

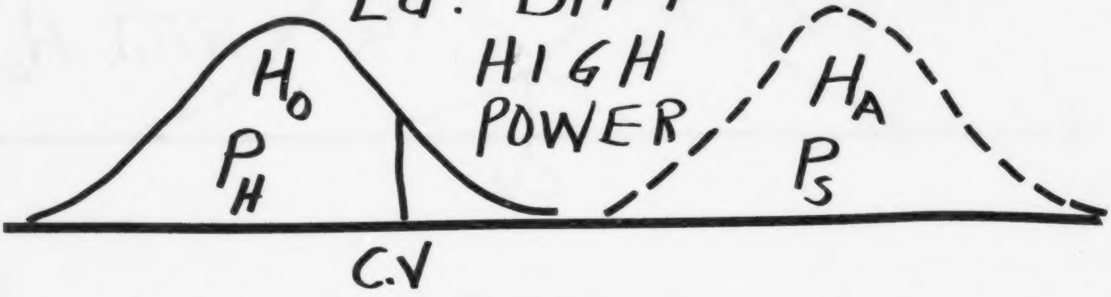
SAMPLE SIZE>>n

HIGH POWER; REJ. H_0

α THE SAME

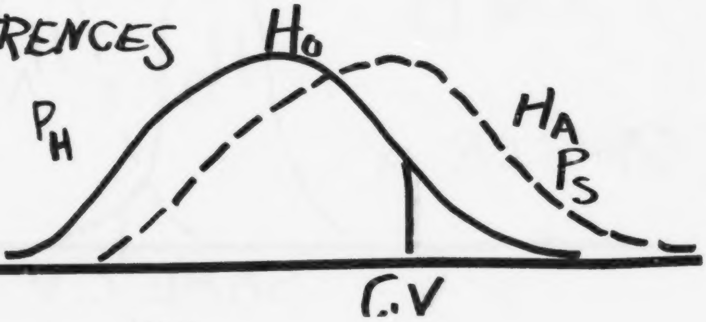

H_0
H_A
HIGH PWR
C.V.

LG. DIFF


H_0
P_H
HIGH POWER
H_A
P_S
C.V

SMALL DIFFERENCES

LOW POWER

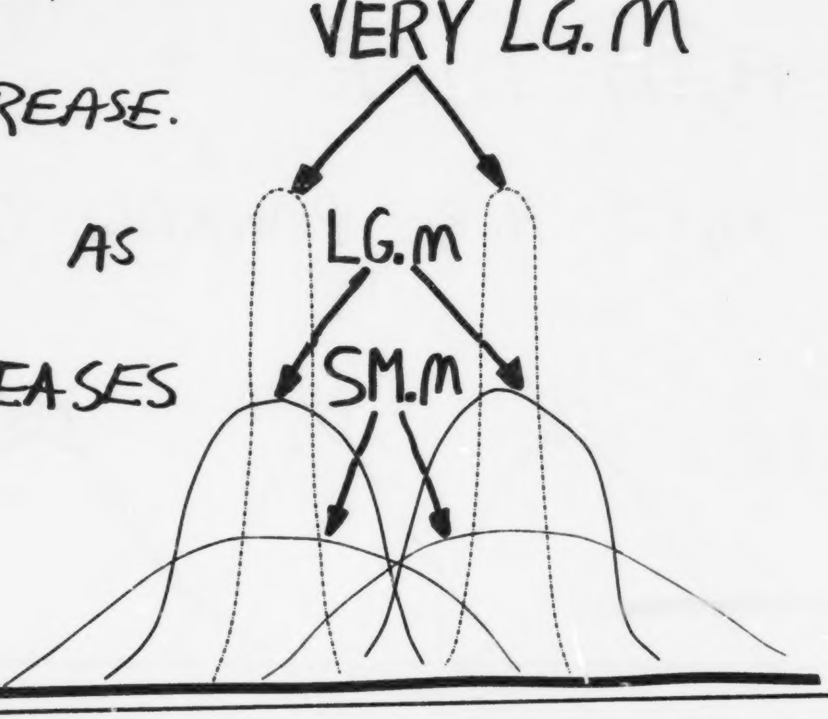

H_0
P_H
H_A
P_S
C.V

CONSTANT

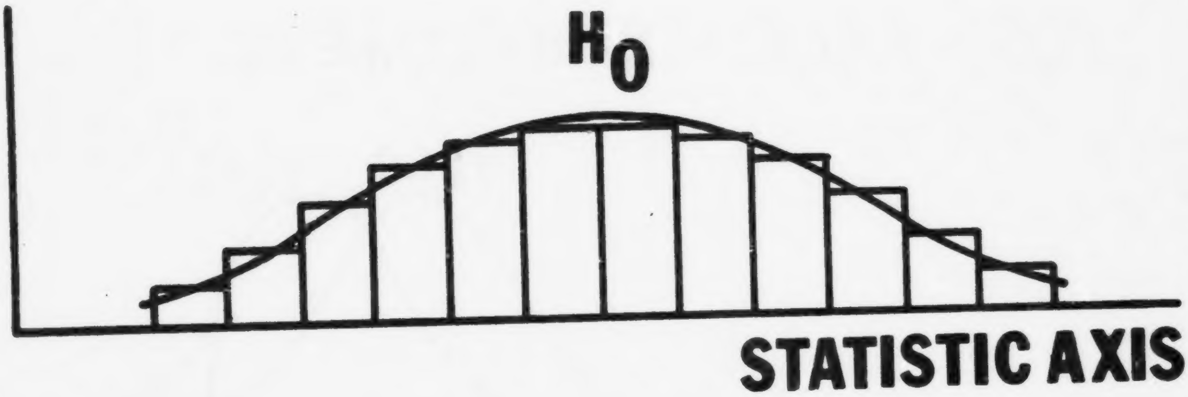
CV → DECREASE.

α & β AS

n INCREASES


VERY LG. n
LG. n
SM. n

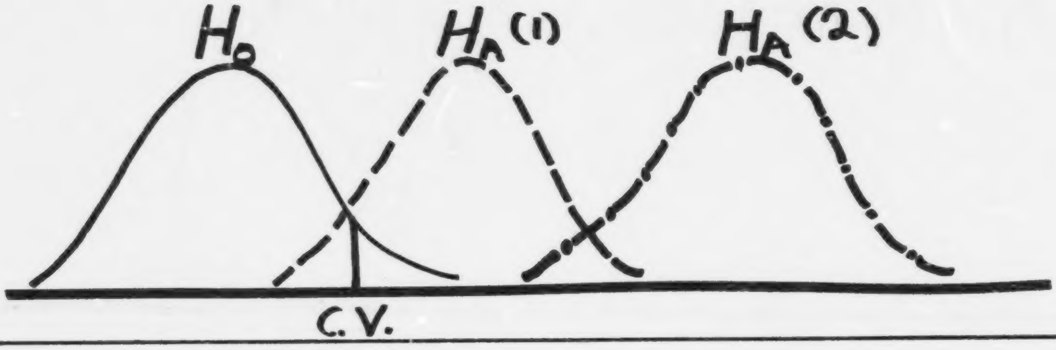
Distribution of the Statistic When the Null Hypothesis Is True



$H_A(1)$ FIXED PARA. VALUES

$H_A(2)$ DIFF. " "

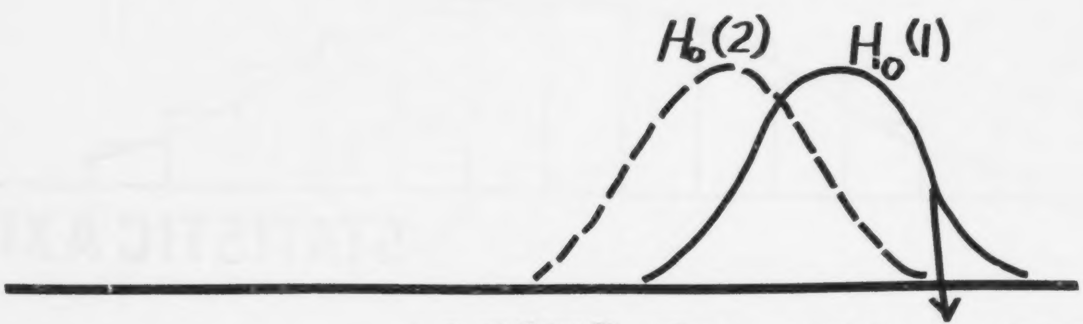
$H_A(2)$ MORE POWERFULL THAN $H_A(1)$


H_0
$H_A(1)$
$H_A(2)$
C.V.

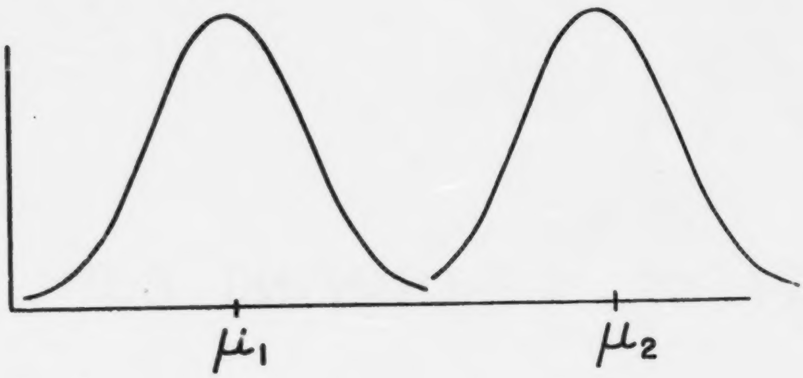
$H_0(1)$ MEETS TESTS ASSUMPTIONS

$H_0(2)$ FAILS " "

$H_0(2)$ LACKS ROBUSTNESS



CORRECT C.V. FROM TABLES

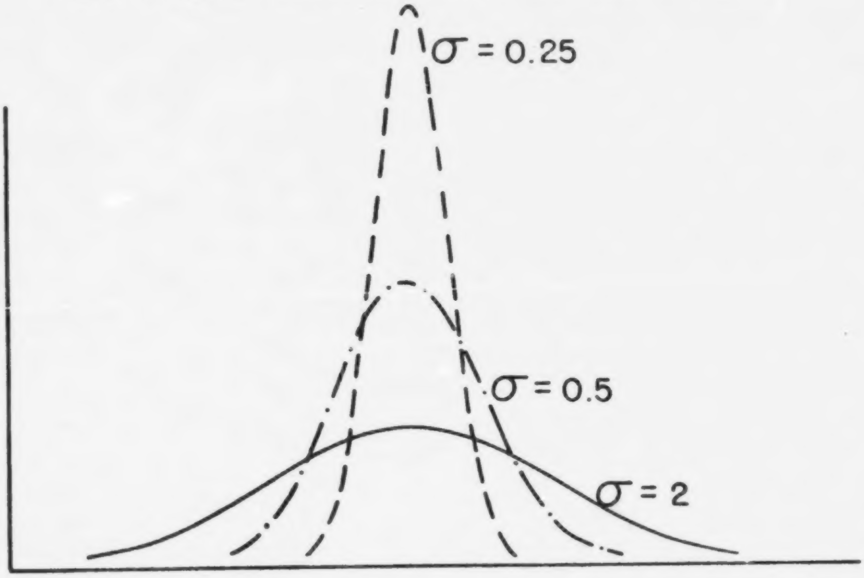


EFFECT OF A CHANGE IN μ

μ_2 LARGER THAN μ_1

σ SAME FOR THE CURVES.

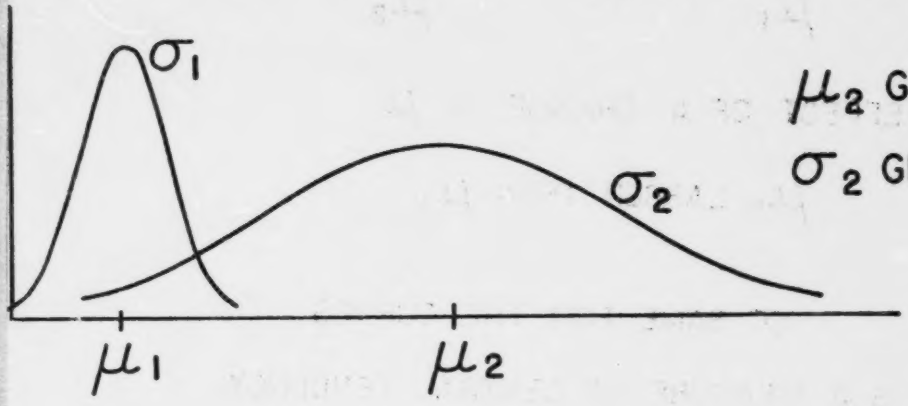
μ IS A MEASURE OF CENTRAL TENDENCY OR WHERE VALUES TEND TO CENTRALIZE.


$\sigma = 0.25$
$\sigma = 0.5$
$\sigma = 2$

EFFECT OF CHANGING σ ALONE AND KEEPING μ THE SAME.

σ IS A MEASURE OF SPREAD OR VARIATION IN ORIGINAL SCALE UNITS.

EFFECTS OF CHANGING BOTH μ & σ

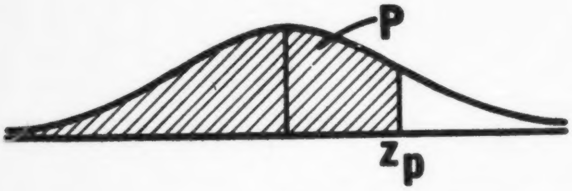


μ_2 GREATER THAN μ_1

σ_2 GREATER THAN σ_1

CUMULATIVE NORMAL DISTRIBUTION–VALUES OF P

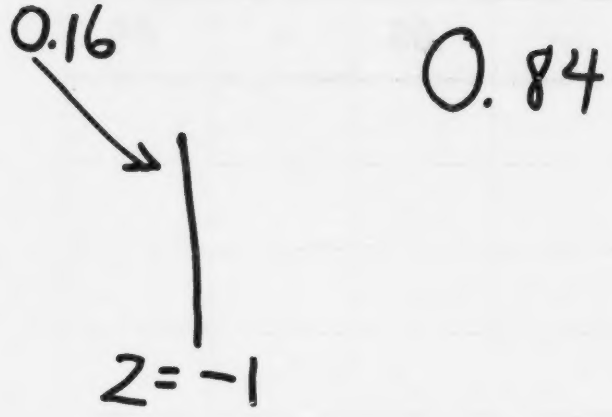
Table A-1
T-2

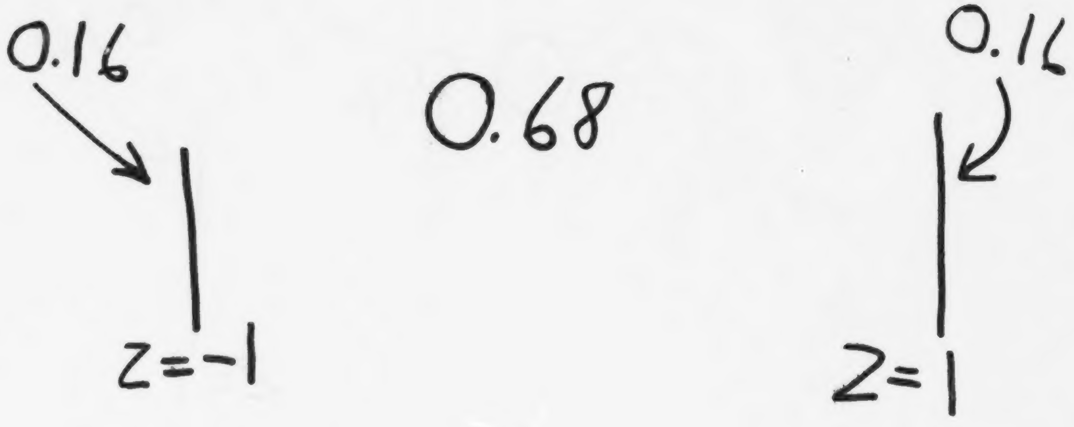


z_p	.00	.01	⋯	.05	⋯	.09
.0						
⋮						
1.0	.8413					
⋮						
2.0				.9798		

NoR. DITB

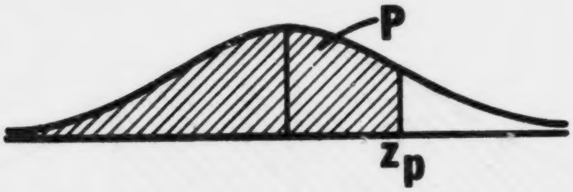
0.84


0.16
0.84
Z=-1

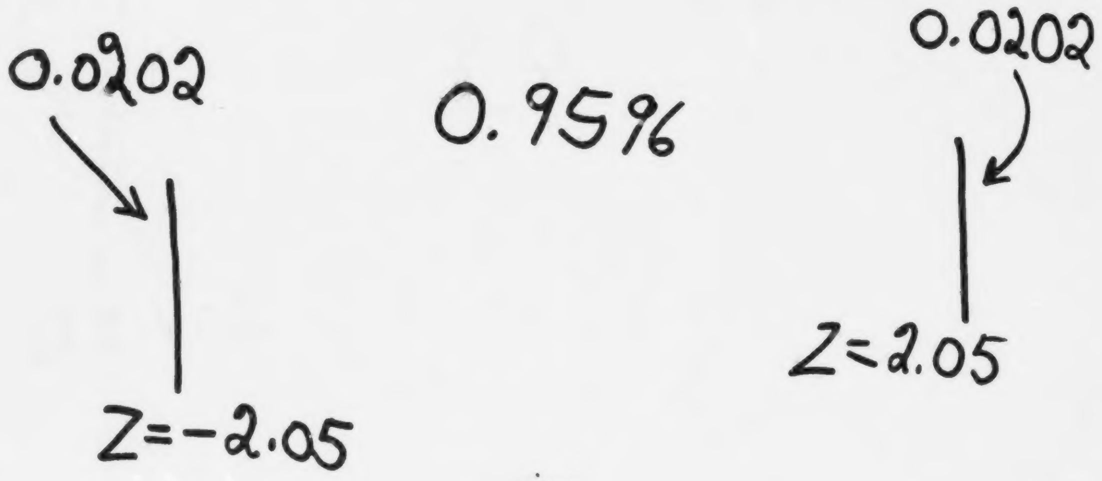

0.16
0.68
0.16
Z=-1
Z=1

CUMULATIVE NORMAL DISTRIBUTION–VALUES OF P

Table A-1
T-2

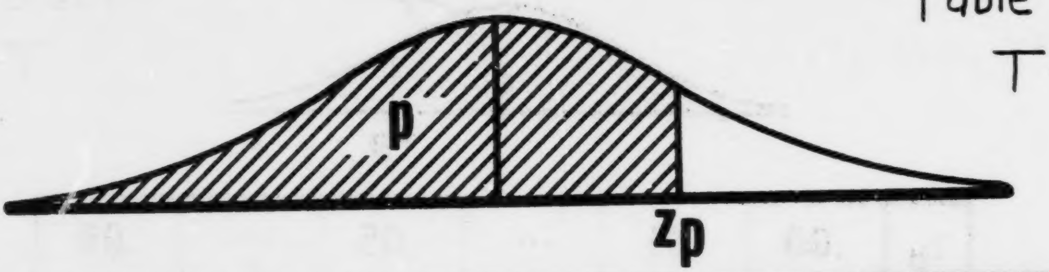


z_p	.00	.01	···	.05	···	.09
.0						
⋮						
1.0	.8413					
⋮						
2.0				.9798		

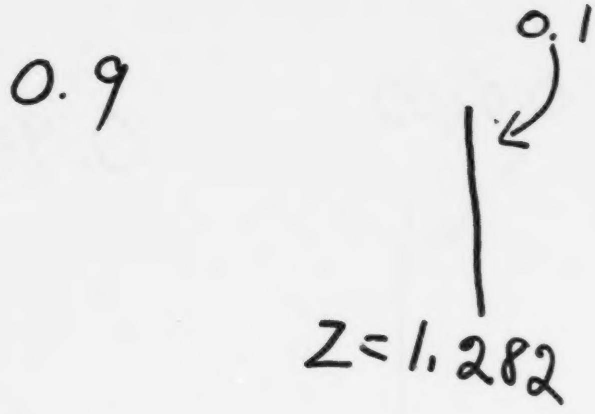

0.0202
0.9596
0.0202
Z=−2.05
Z=2.05

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090



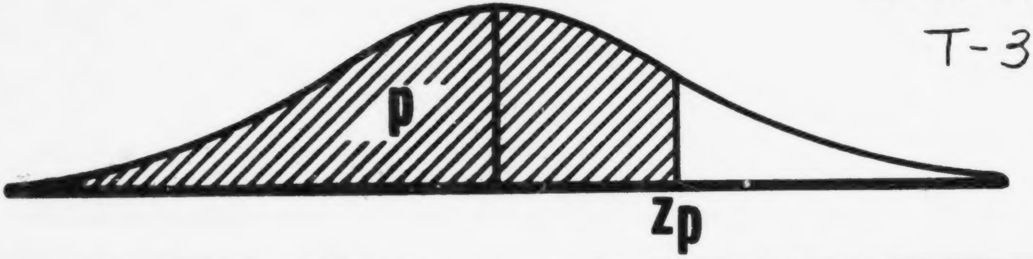
0.05

0.95

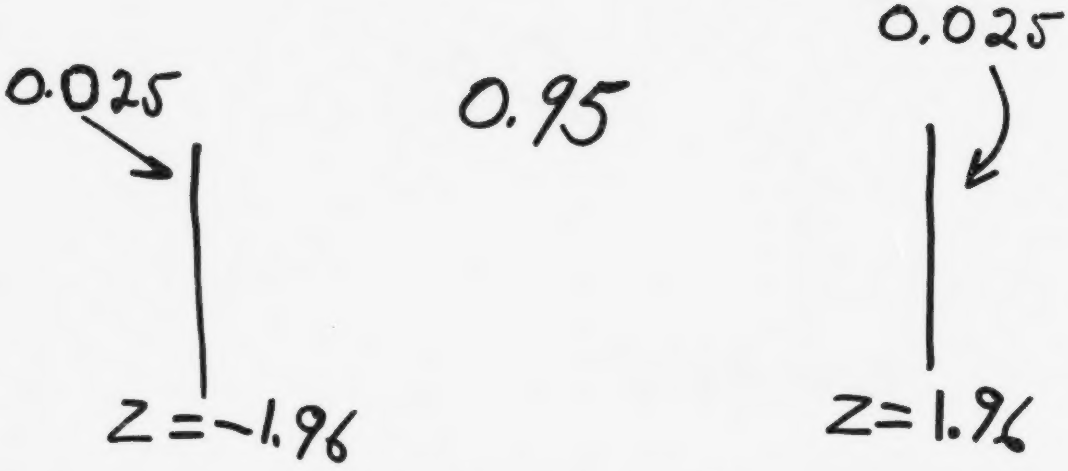
z = -1.645

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3

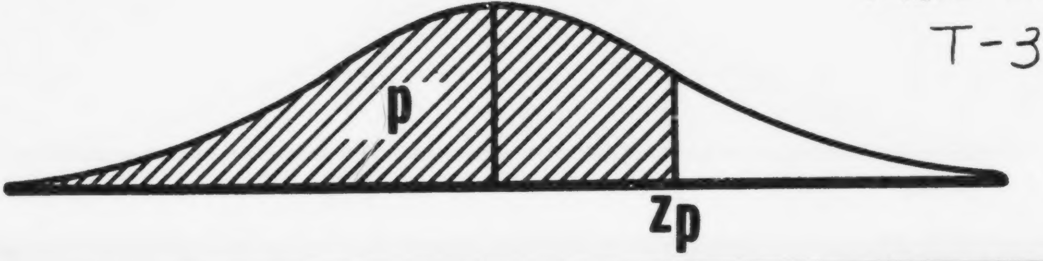


P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090


0.025
0.95
0.025
Z=-1.96
Z=1.96

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090

TABLE A-1 & A-2 FOR

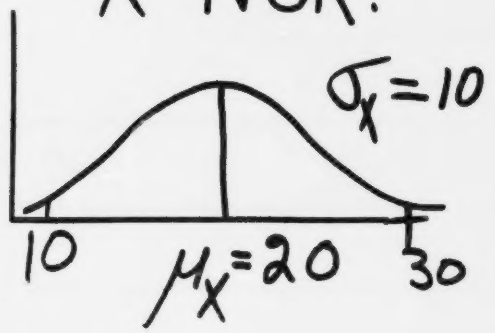
UNIT NORMAL | $\mu = 0$ | $\sigma = 1$

X = ANY NOR. | ME. $= \mu_X$ | S.D. $= \sigma_X$

TRANS. $Z = \dfrac{X - \mu_X}{\sigma_X}$ STAD. VAR.

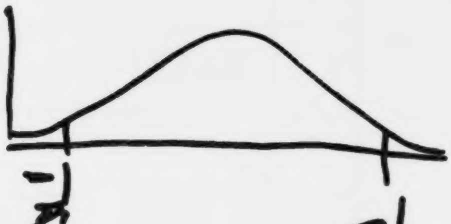
UNIT NOR.

X–NOR.


$\sigma_X = 10$
10
$\mu_X = 20$
30

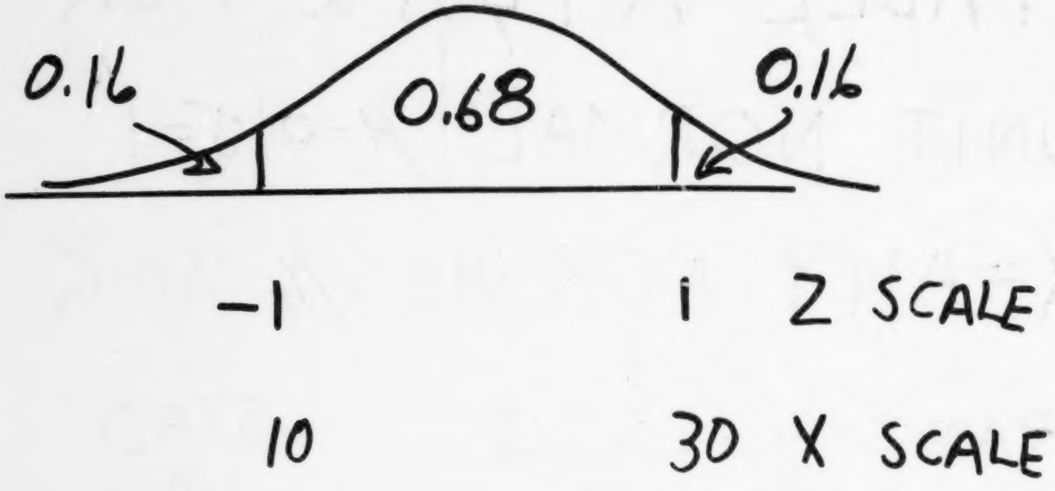
FIND AREA UNDER CURVE

$X = 10$ TO $X = 30$



$Z = (10-20)/10 = -1$

$Z = (30-20)/10 = 1$


0.16
0.68
0.16
−1
1
Z SCALE
10
30
X SCALE

CUMULATIVE NORMAL DISTRIBUTION–VALUES OF P

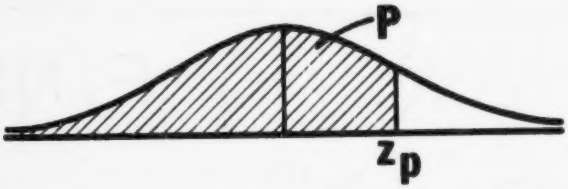


Table A-1
T-2

z_p	.00	.01	···	.05	···	.09
.0						
⋮						
1.0	.8413					
⋮						
2.0				.9798		

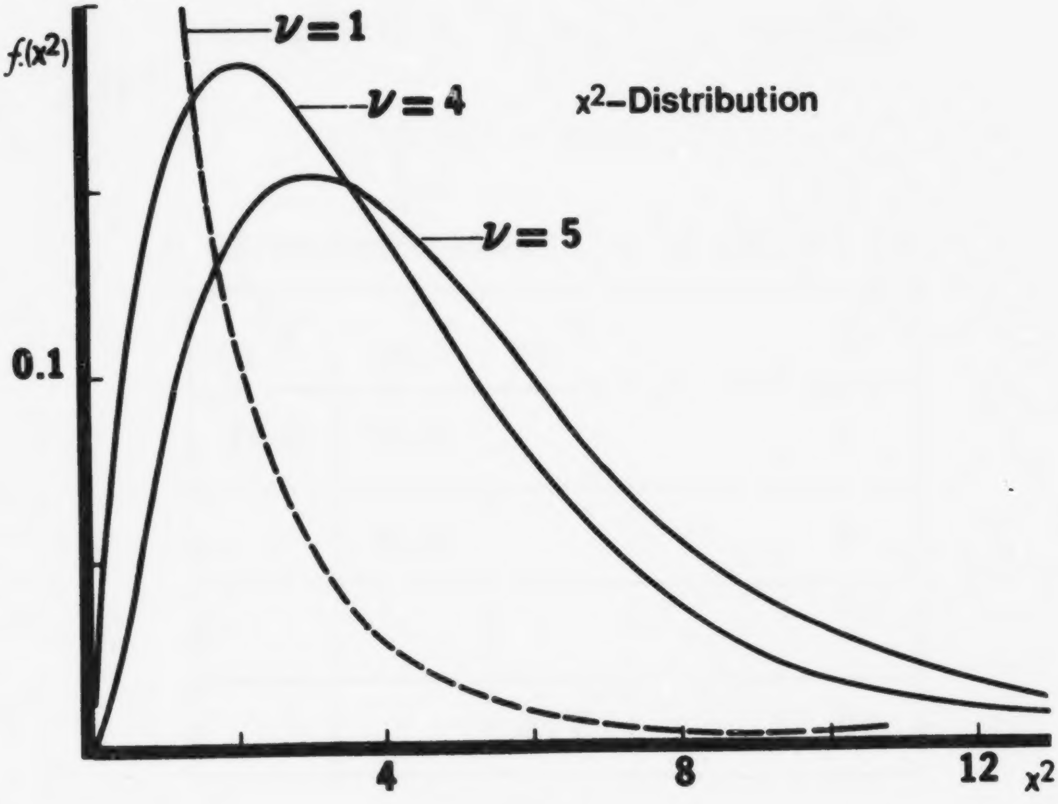
$Z \in$ NOR. DISTB. ⋮ MEAN=0 ⋮ ST.DV.=1

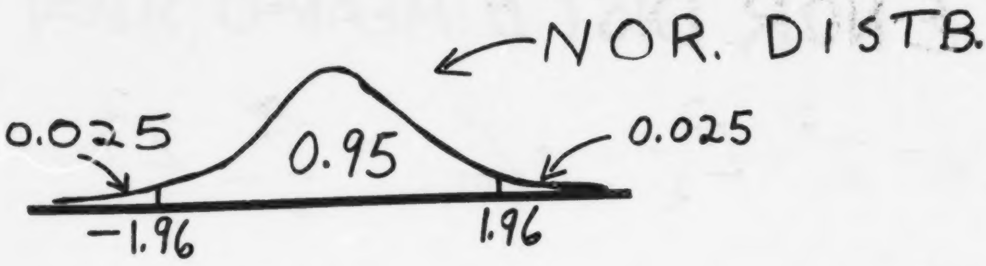
$$\text{R.V.} = Z_1^2 + Z_2^2 + \cdots + Z_n^2 = \sum_{i=1}^{n} Z_i^2$$

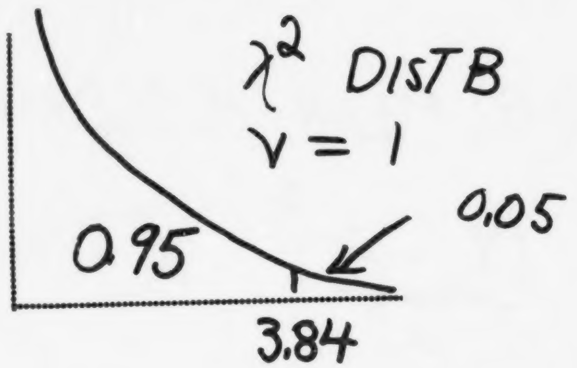
R.V. DISTB AS χ^2 (df OR $\nu = n$)

$X \in$ NOR. DISTB. ⋮ MEAN=μ ⋮ ST.DV.=σ

$$\text{R.V.} = \sum_{i=1}^{n} \frac{(X_i - \mu)^2}{\sigma^2} \quad \chi^2 (\nu = n)$$




NOR. DISTB.
0.025
0.95
0.025
−1.96
1.96


χ^2 DISTB
$\nu = 1$
0.95
0.05
3.84

1.96
1.96
1176
1764
196
3.8416

PERCENTILES OF THE χ^2 DISTRIBUTION

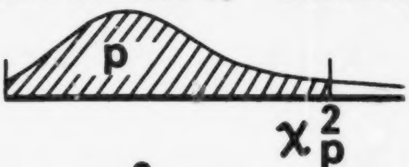
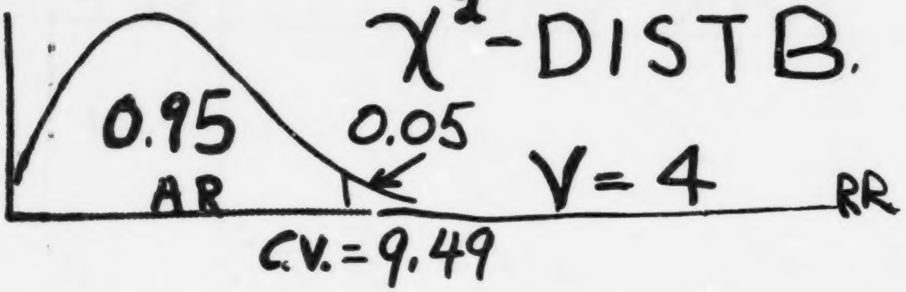


Table A-3
T-4

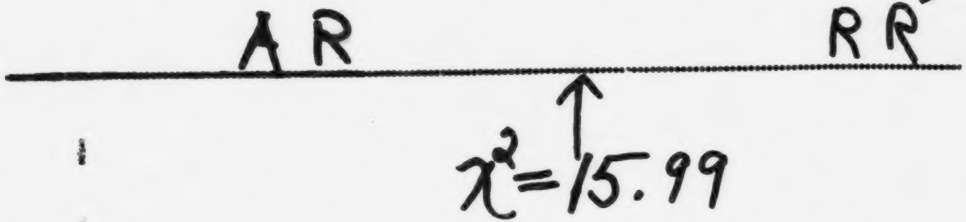
THE VALUES OF χ^2_p CORRESPONDING TO p

df	. . .	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.99}$
1			3.84	6.63
4			9.49	
⋮				
10		15.99		

χ^2-DISTB.


0.95
AR
0.05
$\nu = 4$
RR
C.V. = 9.49

χ^2 DISTB. $\nu = 10$ $\alpha = 0.1$


AR
RR
$\chi^2 = 15.99$

PERCENTILES OF THE χ^2 DISTRIBUTION

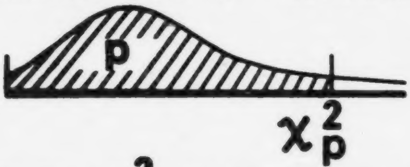


Table A-3
T-4

THE VALUES OF χ^2_p CORRESPONDING TO p

df	...	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.99}$
1			3.84	6.63
4			9.49	
⋮				
10		15.99		

RAN. VAR.	df OR ν
Z^2	1
$Z_1^2 + Z_2^2 + \cdots + Z_n^2 = \sum_{i=1}^{n} Z_i^2$	n
$\sum_{i=1}^{n} (X_i - \mu)^2 / \sigma^2$	n
$\sum_{i=1}^{n} \frac{(X_i - \bar{X})^2}{\sigma^2}$	$n-1$

X ε NOR. DISTB MEAN=μ S.DIV.=σ

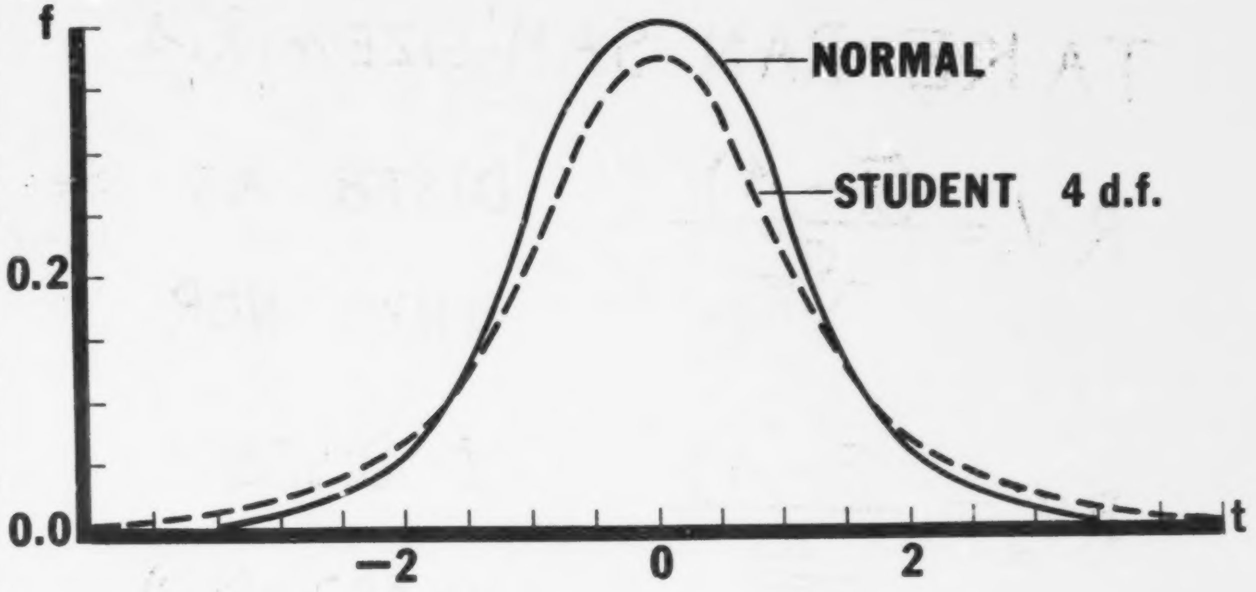
TAKE RAN SAM SIZE=n $\bar{X}$ s

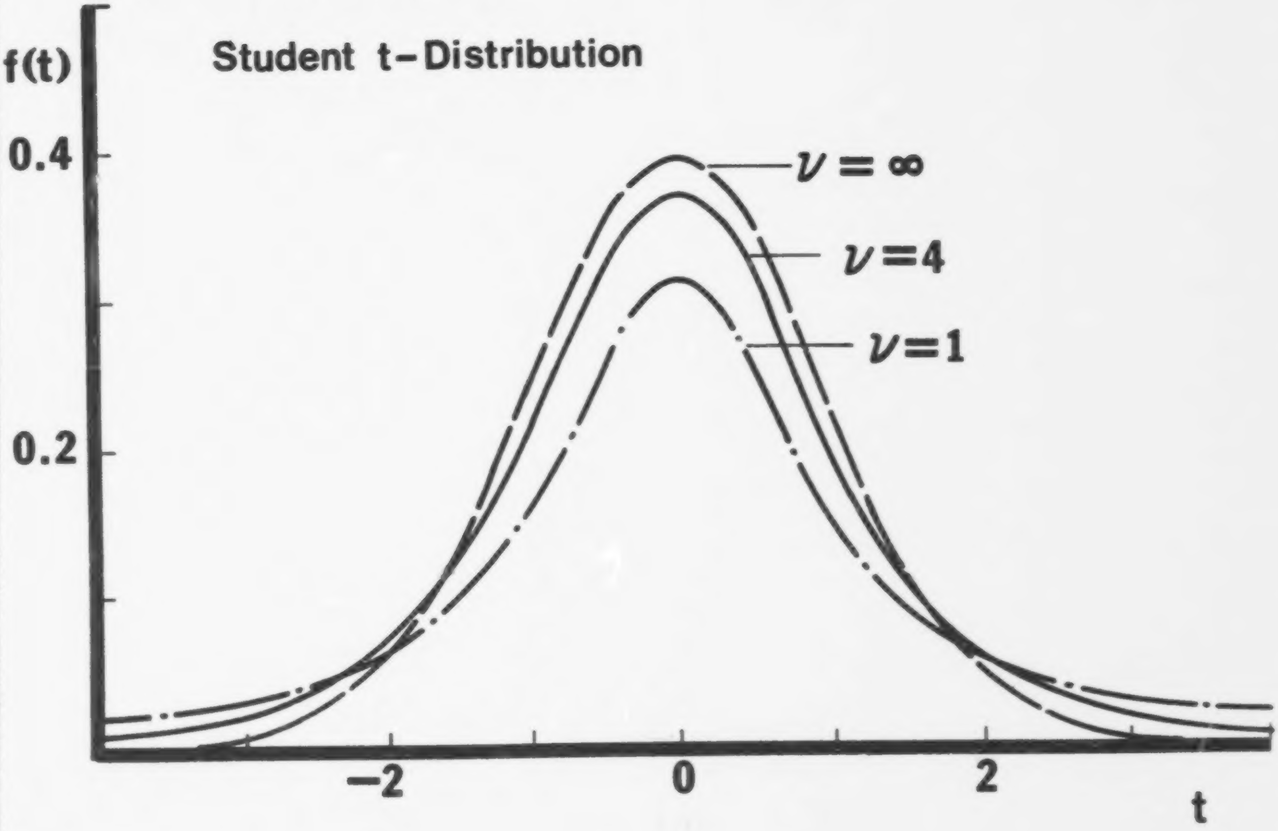
$RV_1 = \dfrac{(\bar{X}-\mu)}{\frac{\sigma}{\sqrt{n}}}$ DISTB. AS UNIT NOR. See 1B-9

$RV_2 = \dfrac{(\bar{X}-\mu)}{\frac{s}{\sqrt{n}}}$ t- DISTB.

df OR $\nu=(n-1)$

Normal vs Student Distribution




Student t-Distribution

LET $n(\bar{X})$ = SAM. SIZE FOR $\bar{X}$

LET $n(s)$ = " " " s

R. V. = $\dfrac{\bar{X} - \mu}{\frac{s}{\sqrt{n(\bar{X})}}}$ $\quad \nu = n(s) - 1$

= df FOR s

IF $n(\bar{X}) = n(s) = n$ ⁞ $\nu = n - 1$

t-DISTB. V=24

0.95

0.05

AR

RR

C.V=1.711

Percentiles of the t Distribution

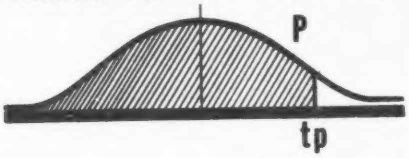
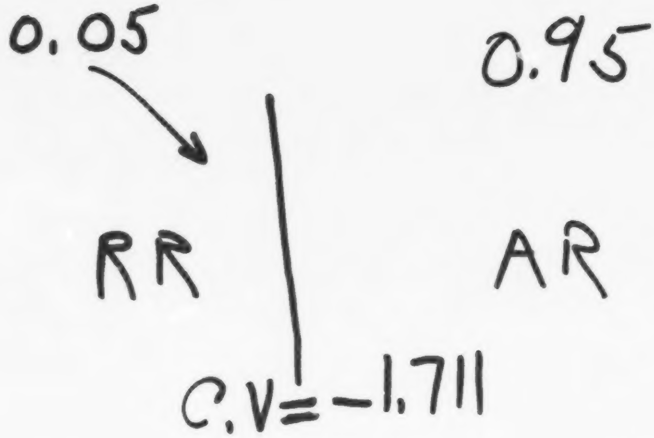


Table A-4
T-5

df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671		

$t \qquad \nu = 24$


0.05
0.95
RR
AR
C.V = -1.711

Percentiles of the t Distribution

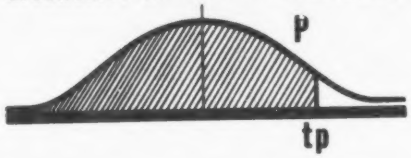
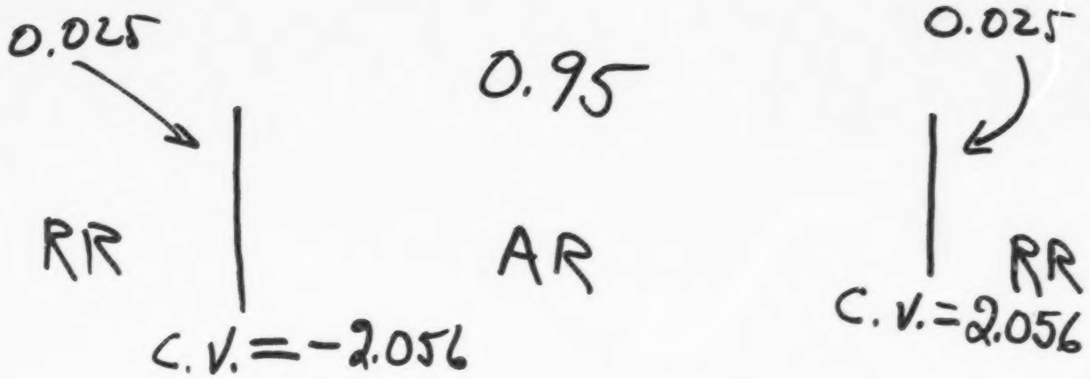


Table A-4
T-5

df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671		

t $\quad V=26$


0.025
0.95
0.025
RR
AR
RR
C.V.=-2.056
C.V.=2.056

Percentiles of the t Distribution

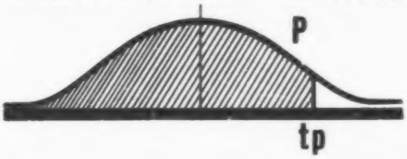
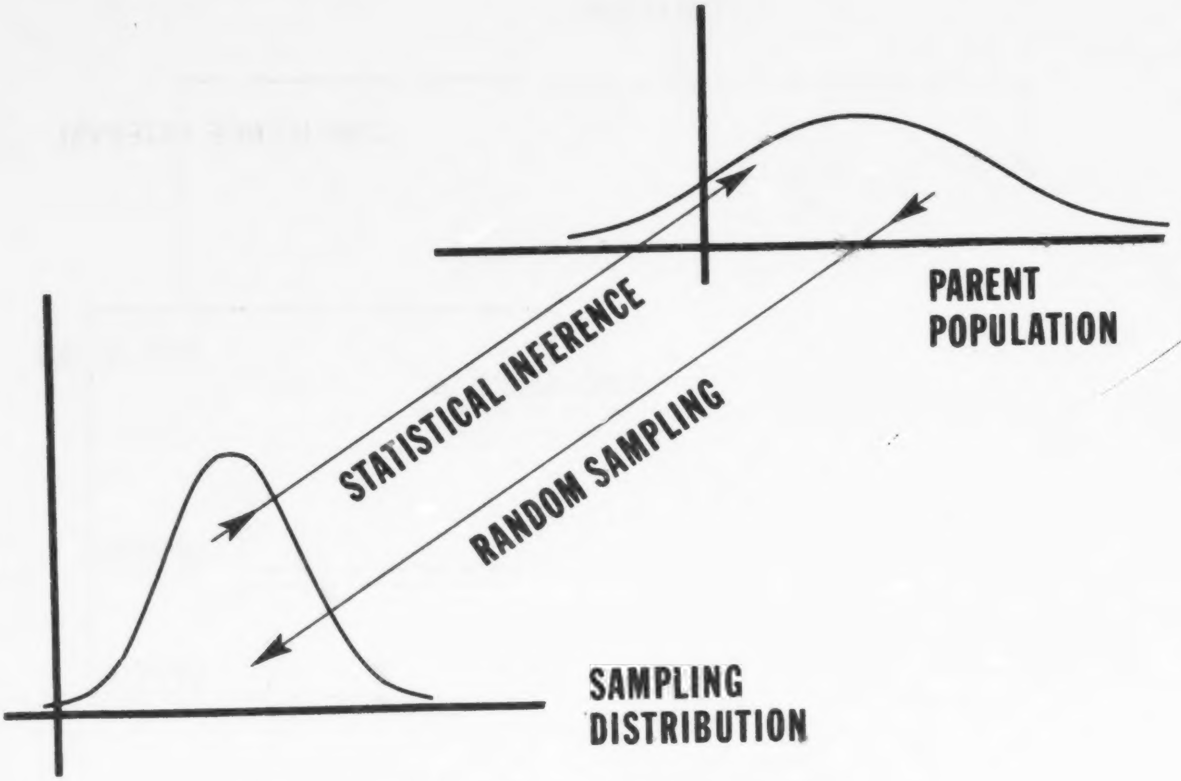
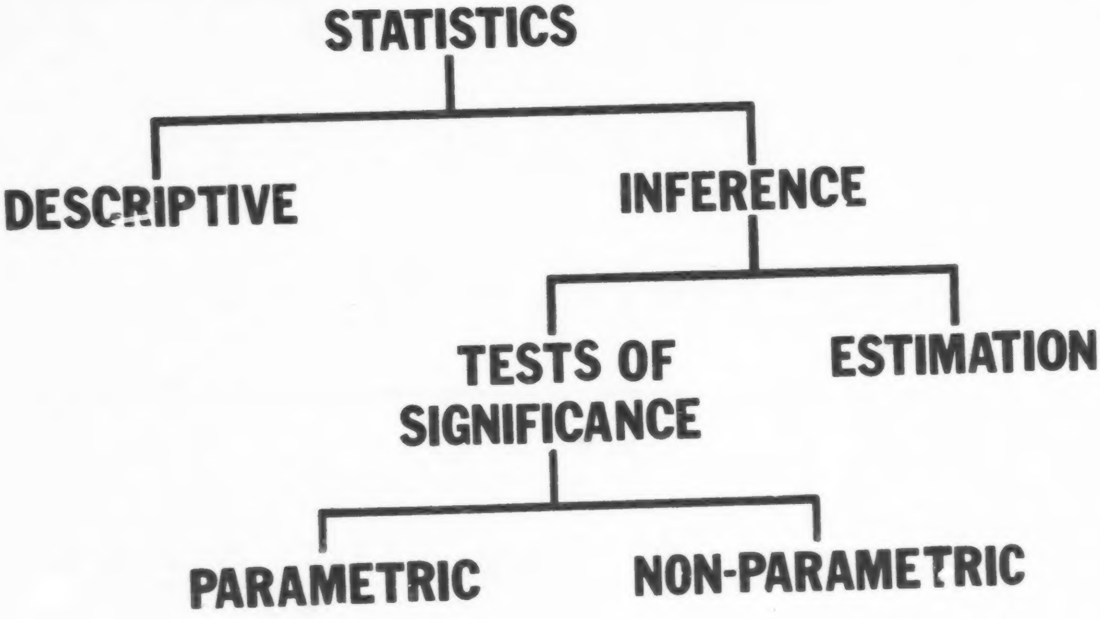


Table A-4
T-5

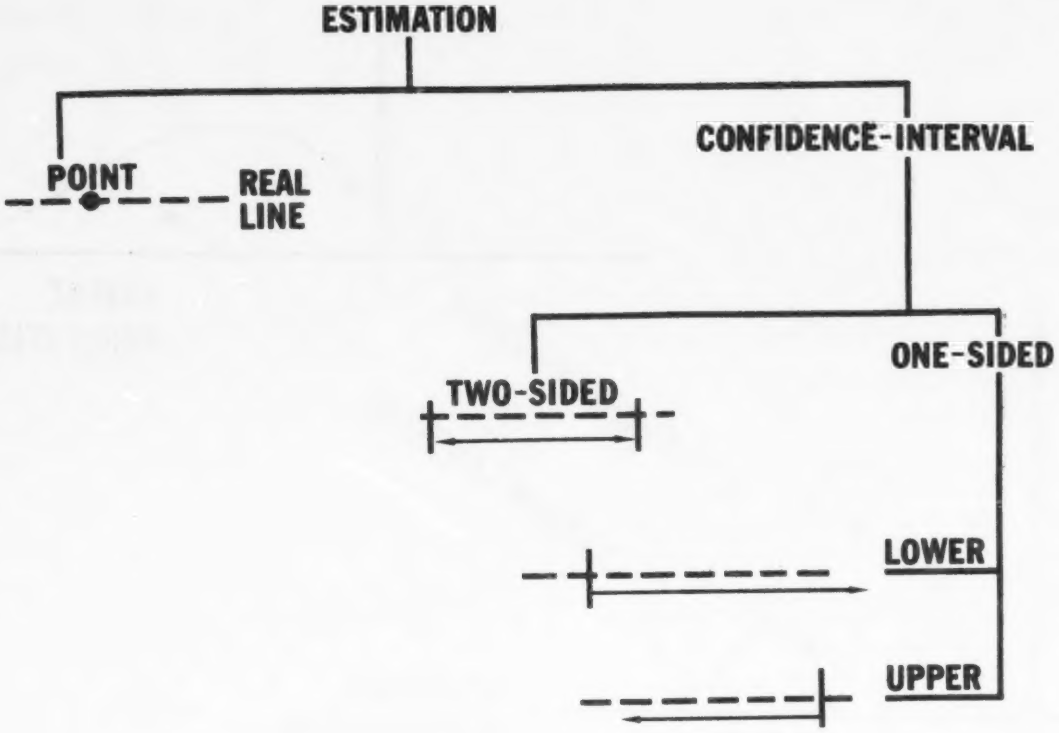
df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671		


STATISTICAL INFERENCE
RANDOM SAMPLING
PARENT POPULATION
SAMPLING DISTRIBUTION

Classification of Statistic



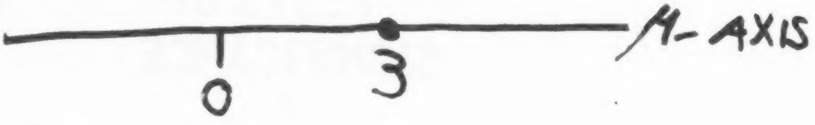
- STATISTICS
  - DESCRIPTIVE
  - INFERENCE
    - TESTS OF SIGNIFICANCE
      - PARAMETRIC
      - NON-PARAMETRIC
    - ESTIMATION


ESTIMATION
POINT
REAL LINE
CONFIDENCE-INTERVAL
TWO-SIDED
ONE-SIDED
LOWER
UPPER

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \sum_{i=1}^{n} X_i/n$$

$$X_1 = 2 \;\vdots\; X_2 = 3 \;\vdots\; X_3 = 4 \;\vdots\; n = 3$$

$$\bar{X} = (2+3+4)/3 = 9/3 = 3$$



$\bar{X}$ EST. μ PAR. POP. MEAN

$$s^2 = \frac{(X_1-\bar{X})^2+(X_2-\bar{X})^2+\cdots+(X_n-\bar{X})^2}{n-1} = \sum_{i=1}^{n}\frac{(X_i-\bar{X})^2}{n-1}$$

$$s^2=\left[(2-3)^2+(3-3)^2+(4-3)^2\right]/2=\left[1+0+1\right]/2=1$$

$$s^2=\left[n\sum X^2-(\sum X)^2\right]/(n-1)n=\left[3(4+9+16)-9^2\right]/2(3)$$

$$=(87-81)/6=6/6=1$$

s^2 EST. OF σ^2 PAR. POP. VAR.

$s_B = \sqrt{s^2}$ BIASED EST. OF σ
PAR. POP. STAND. DEV.

$s_B = \sqrt{1} = 1$; IF $n > 30$; BIAS $< 1\%$

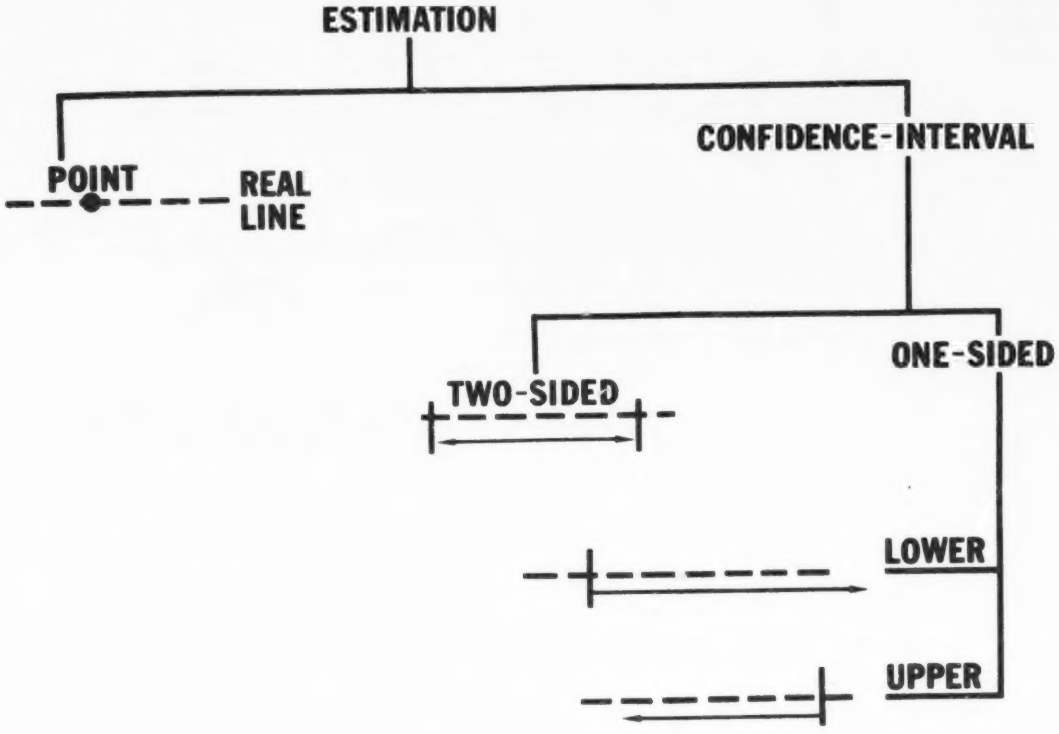
$s_U = s_B / (\text{COR. FAC.}) = 1/0.886 = 1.13$

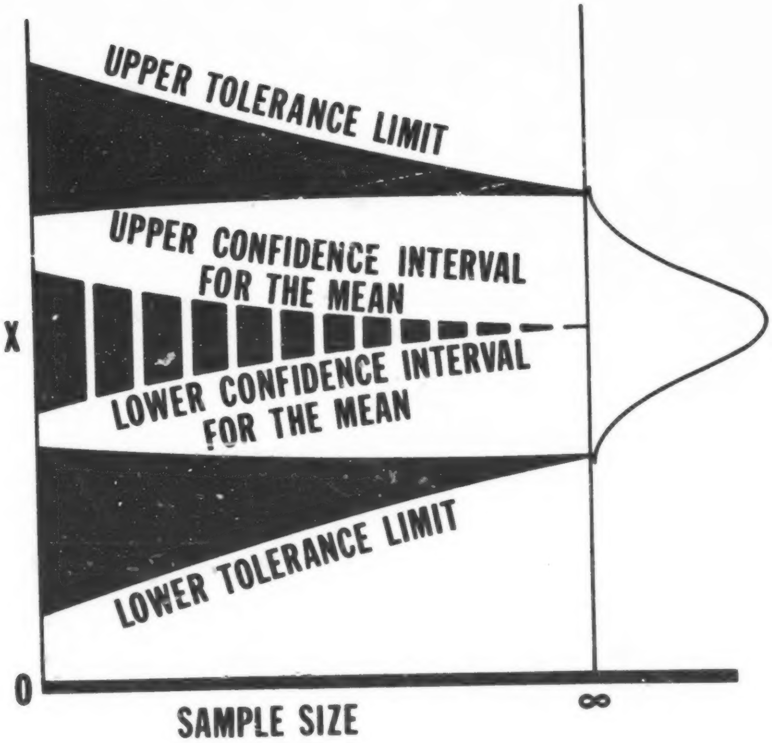
IF $n \geq 6$; $s_U \cong \sqrt{\frac{n-1}{n-1.5}}\, s_B$

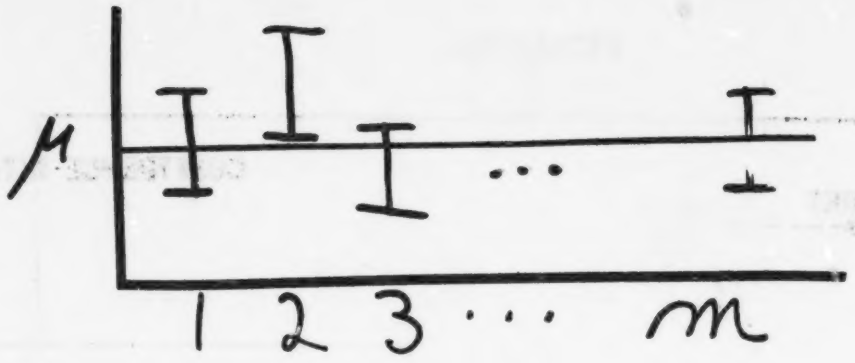
$s_U = \sqrt{(3-1)/(3-1.5)}\,(1) = 1.15$

Page 1-10

SAMPLE SIZE, n	s, IS AN UNBIASED ESTIMATE OF:
2	0.797σ
3	0.886
4	
⋮	


ESTIMATION
POINT
REAL LINE
CONFIDENCE-INTERVAL
TWO-SIDED
ONE-SIDED
LOWER
UPPER


UPPER TOLERANCE LIMIT
UPPER CONFIDENCE INTERVAL FOR THE MEAN
X
LOWER CONFIDENCE INTERVAL FOR THE MEAN
LOWER TOLERANCE LIMIT
0
SAMPLE SIZE
∞



DIFF. ESTIMATES

SUCCESS/$m \longrightarrow 0.95$ AS

$m \longrightarrow \infty$

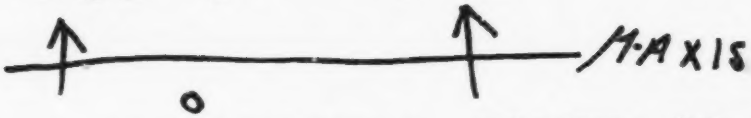
100(1−α)% CONF. INT. FOR μ(2-S)

σ (KWN) = 5 ⋮ $\alpha = 0.05$ ⋮ $n = 3$ ⋮ $\bar{X} = 3$

$$\bar{X} - Z_{1-0.5\alpha}\frac{\sigma}{\sqrt{n}} < \mu < \bar{X} + Z_{1-0.5\alpha}\frac{\sigma}{\sqrt{n}}$$

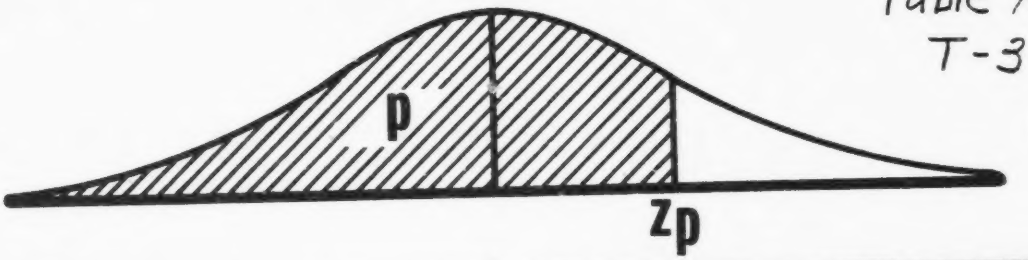
$$3 - 1.96(5)/\sqrt{3} < \mu < 3 + 1.96(5)/\sqrt{3}$$

$$-2.66 < \mu < 8.66$$



CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090

100(1-α)% CONF. INT. μ (1-SIDED)

σ KNOWN

U.L. $\mu < \bar{X} + Z_{1-\alpha} \frac{\sigma}{\sqrt{n}}$

L.L. $\underline{\bar{X} - Z_{1-\alpha} \frac{\sigma}{\sqrt{n}}} < \mu$

100(1-α)% CONF. INT. μ (2-SIDED)

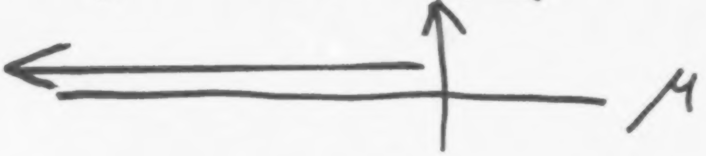
σ UNKNOWN ⋮ s WITH df=ν KNOWN

$$\bar{X} - t_{1-0.5\alpha}(df=\nu) \frac{s}{\sqrt{n}} < \mu < \bar{X} + t \frac{s}{\sqrt{n}}$$

100(1−α)% CONF. INT. μ (1-SIDED)

σ UNKNOWN | s WITH $df=\nu$, KNOWN

U.L. $\mu < \bar{X} + t_{1-\alpha}(df=\nu)\,\frac{s}{\sqrt{n}}$



L.L: $\bar{X} - t_{1-\alpha}(df=\nu)\,\frac{s}{\sqrt{n}} > \mu$

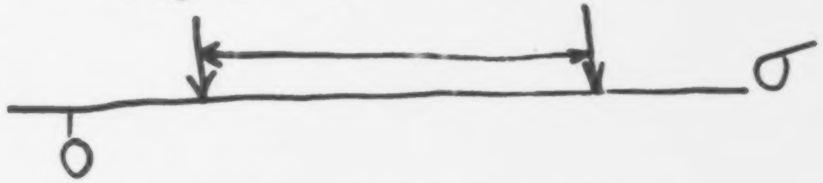
100(1-α)% CONF. INT. σ (2-SIDED)

$\alpha = 0.05$ ¦ $s = 1.13$ ¦ $df = 2$

$$B_L s < \sigma < B_U s$$

$$0.4581(1.13) < \sigma < 4.859(1.13)$$

$$0.52 < \sigma < 5.49$$



FACTORS FOR COMPUTING TWO-SIDED CONFIDENCE LIMITS FOR σ

Table A-20
T-34
&
T-35

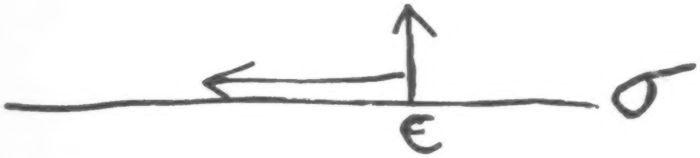
DEG. OF FREEDOM ν	$\alpha = 0.05$	
	B_U	B_L
1	17.79	0.3576
2	4.859	0.4581
3	3.183	0.5178
⋮		
100	1.157	0.8757

ALSO $\alpha = 0.01, 0.001$

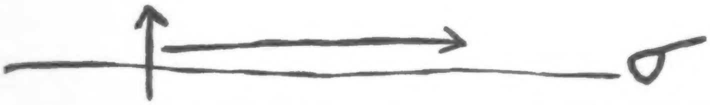
100(1-α)% CONF. INT. σ (1-SIDED)

TABLE A-21

U.L. $\sigma < A_{1-\alpha}\, s$



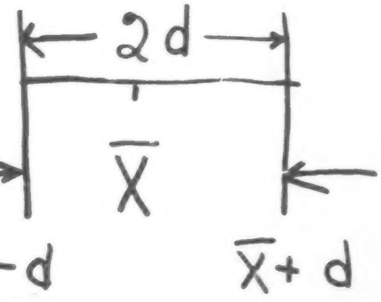
L.L. $A_{\alpha}\, s < \sigma$



---

d≡MARIN OF ERROR

100(1-α)% CONF INT →



σ (KWN) | $\sqrt{n}\,(\bar{x}-\mu)/\sigma = Z$ NOR. DISTB.

$n = (Z_{1-0.5\alpha}\,\sigma/d)^2$ | $n = (t_{1-0.5\alpha}(df=v)\, s/d)^2$

σ (UNK) | s KNOWN WITH df = V

σ & Δ UNK; PICK d, α; GUESS σ_G ; $\alpha = .05$

$n' = (Z_{1-0.5\alpha}\sigma_G/d)^2 = (1.96(5)/1)^2 = 96.04$

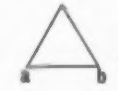
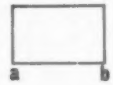
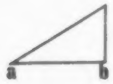
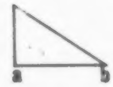
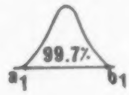
LET n_1 = SAMP. SIZE, FIRST STAGE

IF $n' < 60 \longrightarrow$ LET $n_1 = 0.6n'$

IF $n' \geq 60$ PICK $36 < n_1 << n'$

$36 < n_1 << 97 \longrightarrow$ LET, $n_1 = 61$

Fig. 2-1
Page 2-9

DISTRIBUTION	STANDARD DEVIATION Estimated by
	$\frac{b-a}{4.9}$
	$\frac{b-a}{3.5}$
	$\frac{b-a}{4.2}$
	$\frac{b-a}{4.2}$
	$\frac{b_1-a_1}{6}$

σ & s UNK ⁞ PICK d, α ⁞ GUESS σ_G ⁞ $\alpha = .05$

$n' = (Z_{1-0.5\alpha}\sigma_G/d)^2 = (1.96(5)/1)^2 = 96.04$

LET n_1 = SAMP. SIZE, FIRST STAGE

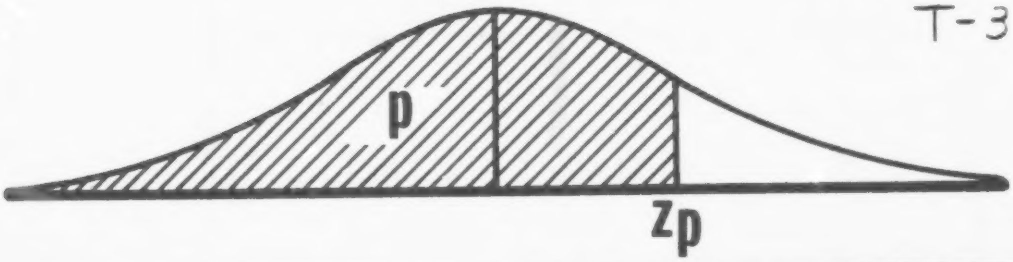
IF $n' < 60 \longrightarrow$ LET $n_1 = 0.6n'$

IF $n' \geq 60$ PICK $36 < n_1 << n'$

$36 < n_1 << 97 \longrightarrow$ LET $n_1 = 61$

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090

$\alpha = 0.05$ ⁞ $s_{61} = 5.0$ ⁞ $df = 60$ ⁞ $d = 1$ ⁞ n_T = TOTAL n

$$n_T = \left(t_{1-0.5\alpha}(\gamma)\, s_{61}/d\right)^2 = \left(2(5)/1\right)^2 = 100$$

n_2 (SECOND STAGE) $= n_T - n_1 = 100 - 61 = 39$

EST. INT. WITH $n=100$

Percentiles of the t Distribution

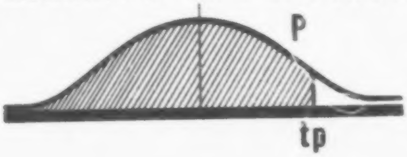


Table A-4
T-5

df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671	2.0	

$P \equiv$ % ERROR FROM TRUE $\sigma = 25\%$

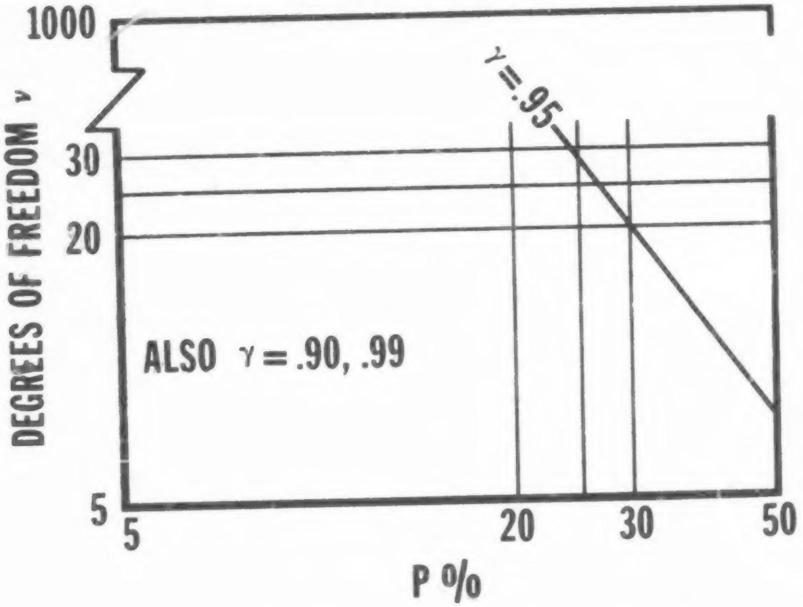
$$\underbrace{\sigma - P\sigma < \sigma < \sigma + P\sigma}_{2P\sigma}$$

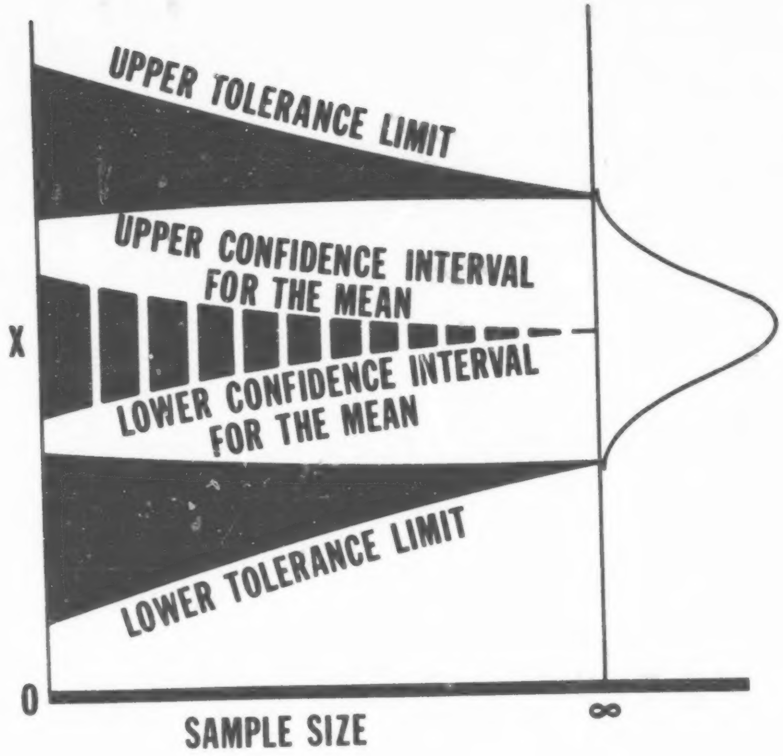
$\gamma \equiv$ CONF. COEFF. $= 0.95$

$df = 29 = n - 1 \longrightarrow n = 30$

NUMBER OF MEASUREMENTS REQUIRED
TO ESTABLISH THE VARIABILITY WITH STATED PRECISION

Fig. 2-2
Page 2-12

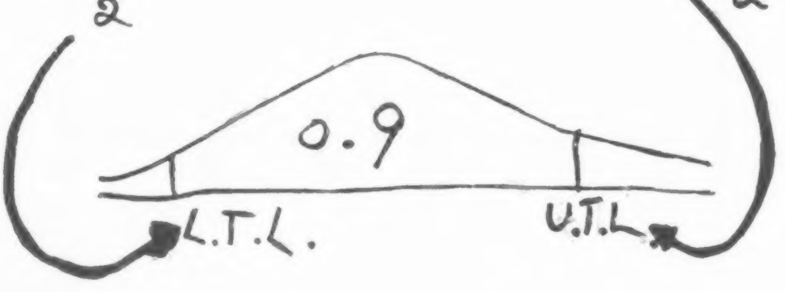



UPPER TOLERANCE LIMIT
UPPER CONFIDENCE INTERVAL FOR THE MEAN
X
LOWER CONFIDENCE INTERVAL FOR THE MEAN
LOWER TOLERANCE LIMIT
0
SAMPLE SIZE
∞

μ AND σ KNOWN

0.9 = P ≡ PROPORTION OF POPULATION BETWEEN LIMITS

$$\mu - Z_{\frac{1+p}{2}}\sigma < P < \mu + Z_{\frac{1+p}{2}}\sigma$$



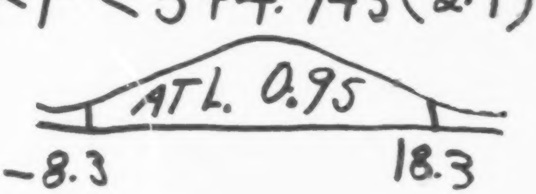
100γ% 2-SIDED TOL. LIMITS FOR
100P% OF POP. VALUES σ & μ-UNK

$\gamma = 0.9$; $P = 0.95$; $\bar{X} = 5$; $s = 2.7$; $n = 4$

$$\bar{X} - Ks < P < \bar{X} + Ks$$

$$5 - 4.943(2.7) < P < 5 + 4.943(2.7)$$

$$-8.3 < P < 18.3$$



FACTORS FOR TWO-SIDED TOLERANCE LIMITS FOR NORMAL DISTRIBUTIONS

Table A-6
T-10

	$\gamma = 0.90$				
	P				
n	0.75	0.90	0.95	0.99	0.999
2					
3					
4			4.943		

$\gamma = 0.75, 0.9, 0.95, 0.99$
n = 2 (VARIOUS) ∞

One Sample Problem

TEST	CONDITIONS
$H_0: \mu =$ CONSTANT	1. σ UNKNOWN
$H_A: \mu \neq$ CONSTANT	2. σ KNOWN
$H_0: \mu \leq$ CONSTANT	3. σ UNKNOWN
$H_A: \mu >$ CONSTANT	4. σ KNOWN
$H_0: \mu \geq$ CONSTANT	5. σ UNKNOWN
$H_A: \mu <$ CONSTANT	6. σ KNOWN

$H_0: \mu = 7.35$ ⋮ $H_A: \mu \neq 7.35$ ⋮ $\alpha = 0.05$ ⋮ $\beta = 0.1$

$$st. = \frac{\sqrt{n}\,(\bar{x} - 7.35)}{s}$$

σ (UNK. & EST.) $= 5$ ⁞ $\alpha = .05$ ⁞ $\beta = .1$ ⁞ $1-\beta = .9$

$\delta = |\mu - \text{CONST.}| \equiv$ AVE. DIFF. TO BE DETECT.

$= |\mu - 7.35| = 1$

$$d = \frac{\delta}{\sigma} = \frac{1}{5} = 0.2$$

$n' = 263 \qquad n = n' + 2 = 265$

Sample Sizes Required...

$\alpha = 0.05$

Table A-8
T-16

d	$1-\beta$						
	.5	.6	.7	.8	.9	.95	.99
.2	97				263		
.4	25	31					
.8	7	8					

ALSO $\alpha = 0.01$

σ (UNK. & EST.) $= 5$ ¦ $\alpha = 0.05$ ¦ $\beta = .5 = 1-\beta$

$$d = \frac{|\mu - 7.35|}{\sigma} = \frac{2}{5} = 0.4$$

$n' = 25 \qquad n = n' + 2 = 27$

Sample Sizes Required...

$\alpha = 0.05$

Table A-8
T-16

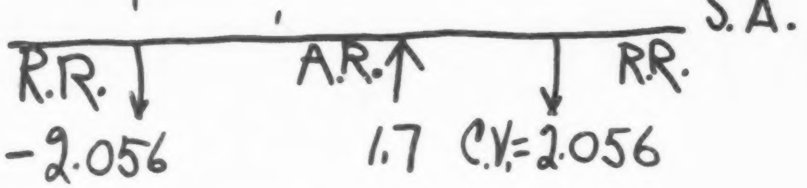
d	$1-\beta$						
	.5	.6	.7	.8	.9	.95	.99
.2	97				263		
.4	25	31					
.8	7	8					

ALSO $\alpha = 0.01$

$H_0: \mu = 7.35$ ⋮ $H_A: \mu \neq 7.35$ ⋮ $\alpha = 0.05$ ⋮ $\beta = 0.5$

$$st. = \frac{\sqrt{n}\,(\bar{x}-7.35)}{s} = \frac{\sqrt{27}\,(9.35-7.35)}{6.1} = 1.7$$

$n = 27$ ⋮ $\bar{x} = 9.35$ ⋮ $s = 6.1$ ⋮ $df = 26$ FOR t-DISTB


S.A.
R.R.
A.R.
R.R.
-2.056
1.7
C.V.=2.056

DEC. DO NOT REJ. H_0 ⋮ CANNOT REJ H_0

Percentiles of the t Distribution

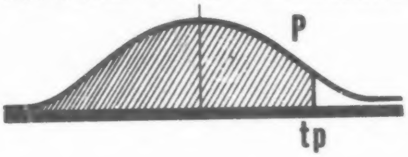


Table A-4

T-5

df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671		

$d(\text{WITH } \sigma = 5) = 0.4 \rightarrow n = 27$

$d(\text{WITH } s = 6.1) = 2/6.1 = 0.32$

$n' < (97+25)/2 = 122/2 = 61$

SAY $n' = 54$ $n = 56$

Sample Sizes Required...

$\alpha = 0.05$

Table A-8

T-16

d	1−β						
	.5	.6	.7	.8	.9	.95	.99
.2	97				263		
.4	25	31					
.8	7	8					

ALSO $\alpha = 0.01$

INTERVAL EST. FOR σ

SMALL CORRECT LARGE

t-TEST

SMALL σ | n | COST

LARGE d | β

CORRECT α

LARGE σ | n | COST

SMALL d | β

CORRECT α

CORRECT σ → CORRECT d | n | α | β | COST

Sample Sizes Required...

$\alpha = 0.05$

d	$1-\beta$						
	.5	.6	.7	.8	.9	.95	.99
.2	97				263		
.4	25	31					
.8	7	8					

Table A-8

T-16

ALSO $\alpha = 0.01$

One Sample Problem

TEST	CONDITIONS
$H_0: \mu =$ CONSTANT	1. σ UNKNOWN
$H_A: \mu \neq$ CONSTANT	2. σ KNOWN
$H_0: \mu \leq$ CONSTANT	3. σ UNKNOWN
$H_A: \mu >$ CONSTANT	4. σ KNOWN
$H_0: \mu \geq$ CONSTANT	5. σ UNKNOWN ←
$H_A: \mu <$ CONSTANT	6. σ KNOWN

$H_0: \mu = 7.35 \;(H_0: \mu \geq 7.35),\; H_A: \mu < 7.35$

$\alpha = 0.05;\; \beta = 0.4$

$$t = \frac{\sqrt{n}\,(7.35 - \bar{X})}{s}$$

$\alpha = 0.05 \;\vert\; \beta = 0.4 \;\vert\; 1-\beta = 0.6$

$$\sigma\,(\text{EST.}) = \frac{b-a}{4.9} = \frac{10.8-1}{4.9} = 2$$

$\delta = (\text{CONST.} - \mu) \equiv$ TRUE DIFF. BTWN. AVES.

$= (7.35 - \mu) = 0.8$

$d = \dfrac{\delta}{\sigma} = \dfrac{0.8}{2} = 0.4 \;\vert\; n' = 23 \;\vert\; n = n' + 2 = 25$

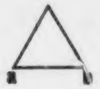
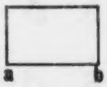
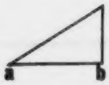
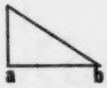
DISTRIBUTION	STANDARD DEVIATION Estimated by
	$\frac{b-a}{4.9}$
	$\frac{b-a}{3.5}$
	$\frac{b-a}{4.2}$
	$\frac{b-a}{4.2}$

Figure 2-1
Page 2-9

Sample Sizes Required...

$\alpha = 0.05$

Table A-9
T-17

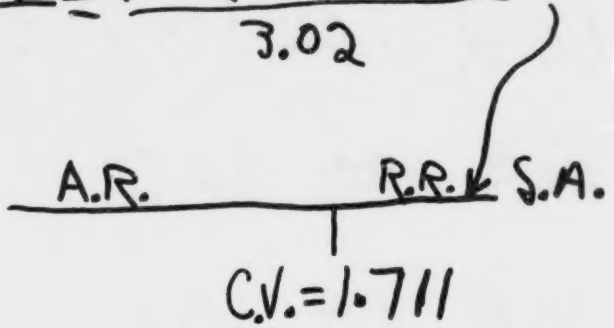
d	1−β						
	.5	.6	.7	.8	.9	.95	.99
.4		23					

$H_0: \mu = 7.35$ ($H_0: \mu \geq 7.35$) $H_A: \mu < 7.35$

$\alpha = 0.05$ $\beta = 0.4$ $n = 25$ $\bar{X} = 6.2$ $s = 3.02$

$$t = \frac{\sqrt{n}\,(7.35-\bar{X})}{s} = \frac{\sqrt{25}\,(7.35-6.2)}{3.02} = 1.9$$

df = 24 t-DISTB.



DEC. REJ. H_0

Percentiles of the t Distribution

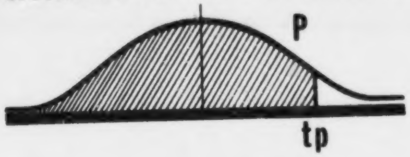


Table A-4

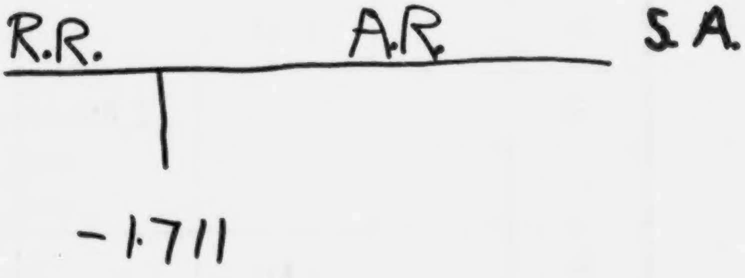
T-5

df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671		

NORMAL DISTB. TEST

σ (INC. & SM.)	σ (INC. & LG.)
LG. d → SM. n	SM. d → LG. n
LG. α C.V. = $f(\sigma)$	SM. α C.V. = $f(\sigma)$
LG. β FOR LG. δ	SM. β FOR LG. δ
SM. β FOR SM. δ	LG. β FOR SM. δ
LOWER COST	HIGHER COST

IF $t = \dfrac{\sqrt{n}\,(\bar{X} - 7.35)}{s}$



Two Sample Problem

TEST	CONDITIONS
$H_0: \mu_A = \mu_B$ $H_A: \mu_A \neq \mu_B$	1. $\sigma_A = \sigma_B$, BOTH UNKNOWN 2. $\sigma_A \neq \sigma_B$, BOTH UNKNOWN 3. σ_A, σ_B BOTH KNOWN ← 4. PAIRED OBSERVATIONS

$H_0: \mu_A = \mu_B$ | $H_A: \mu_A \neq \mu_B$ | $\alpha = 0.05$ | $\beta = 0.4$

$\sigma_A^2 = \sigma_B^2 = 24.5$

$$Z = \frac{|\bar{X}_A - \bar{X}_B|}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}}$$

$\alpha = 0.05$ ¦ $\beta = 0.4$ ¦ $1-\beta = 0.6$ ¦ $n_A = n_B$

$\delta = \mu_A - \mu_B \equiv$ TRUE DIFF. BETWEEN AVES.

$$d = \frac{|\delta|}{\sqrt{\sigma_A^2 + \sigma_B^2}} = \frac{5.6}{\sqrt{24.5 + 24.5}} = \frac{5.6}{\sqrt{49}} = 0.8$$

$n = 8 = n_A = n_B$

Sample Sizes Required...

$\alpha = 0.05$

d	$1-\beta$						
	.5	.6	.7	.8	.9	.95	.99
.2	97				263		
.4	25	31					
.8	7	8					

Table A-8

T-16

ALSO $\alpha = 0.01$

$\alpha \;\vdots\; \beta \;\vdots\; n_A \mp n_B \;\vdots\; \delta \equiv$ SAME

$$d = \frac{|\delta|}{\sqrt{\sigma_A^2 + C\sigma_B^2}}$$ WHERE $\frac{n_A}{n_B} = C$ (CONST.)

$C = 1 \longrightarrow n_A = n_B$

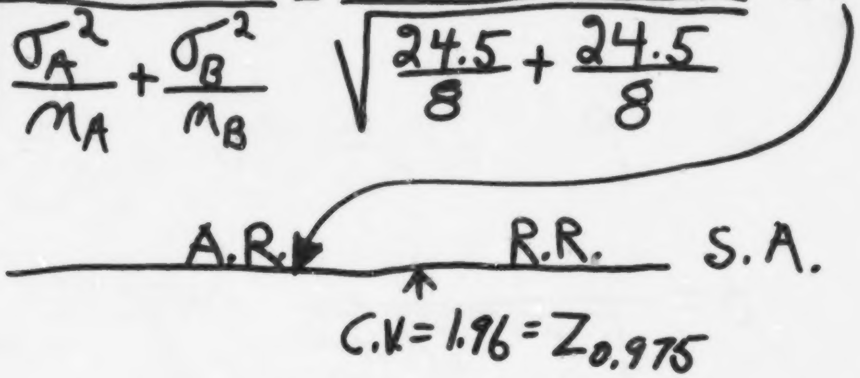
$= 2$ n_A TWICE n_B

$= 1/2$ n_A HALF n_B

$H_0: \mu_A = \mu_B$ | $H_A: \mu_A \neq \mu_B$ | $\alpha = 0.05$ | $\beta = 0.4$

$\sigma_A^2 = \sigma_B^2 = 24.5$ | $\bar{X}_A = 12.1$ | $\bar{X}_B = 16.3$

$$Z = \frac{|\bar{X}_A - \bar{X}_B|}{\sqrt{\frac{\sigma_A^2}{n_A} + \frac{\sigma_B^2}{n_B}}} = \frac{|12.1 - 16.3|}{\sqrt{\frac{24.5}{8} + \frac{24.5}{8}}} = 1.7$$



DEC. DO NOT REJ. H_0

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

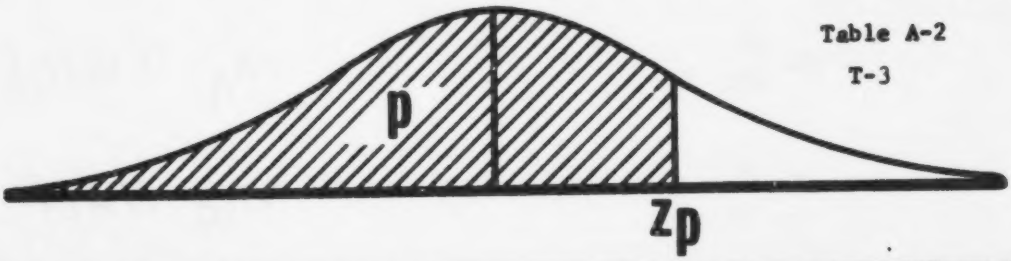


Table A-2
T-3

P	.9	.95	.975	.999
z_p	1.282	1.645	1.960	3.090

Two Sample Problem

TEST	CONDITIONS
$H_0 : \mu_A \leq \mu_B$ $H_A : \mu_A > \mu_B$	5. $\sigma_A = \sigma_B$, BOTH UNKNOWN 6. $\sigma_A \neq \sigma_B$, BOTH UNKNOWN 7. σ_A , σ_B BOTH KNOWN 8. PAIRED OBSERVATIONS

K-SAMPLE PROBLEM

$H_0: \mu_A = \mu_B \, (\mu_A \leq \mu_B) \; \vdots \; H_A: \mu_A > \mu_B$

$\alpha = 0.05 \; \vdots \; \beta = 0.3$

$$t = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{n_A + n_B}{n_A n_B}} \sqrt{\frac{(n_A - 1) s_A^2 + (n_B - 1) s_B^2}{n_A + n_B - 2}}}$$

$\alpha = 0.05$ | $\beta = 0.3$ | $1-\beta = 0.7$ | $n_A = n_B$

$$\sigma_A^2 = \sigma_B^2 = \sigma^2 = 18$$

$\delta = \mu_A - \mu_B =$ TRUE DIFF. BTWN. AVES.

$$d = \frac{\mu_A - \mu_B}{\sqrt{2\sigma^2}} = \frac{2.4}{\sqrt{2(18)}} = \frac{2.4}{6} = 0.4$$

$n' = 30$ $\qquad n = n' + 1 = 31$

Sample Sizes Required...

$\alpha = 0.05$

Table A-9
T-17

d	.5	.6	.7	.8	.9	.95	.99
.1							
.2		91					
.4		23	30				
3.0							

Header over columns .5 to .99: $1-\beta$

ALSO $\alpha = 0.01$

$H_0: \mu_A = \mu_B \; (\mu_A \leq \mu_B) \quad H_A: \mu_A > \mu_B \quad n=31$

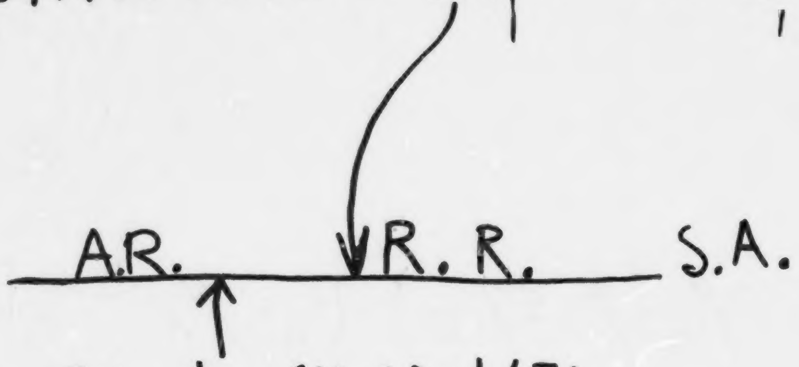
$\alpha = 0.05 \quad \beta = 0.3 \quad \bar{X}_A = 18.2 \quad \bar{X}_B = 16.3 \quad s_A^2 = 18 \quad s_B^2 = 15$

$$t = \frac{\bar{X}_A - \bar{X}_B}{\sqrt{\frac{n_A + n_B}{n_A n_B}} \sqrt{\frac{(n_A - 1) s_A^2 + (n_B - 1) s_B^2}{n_A + n_B - 2}}}$$

$$t = \frac{18.2 - 16.3}{\sqrt{\frac{31+31}{31(31)}} \sqrt{\frac{30(18) + 30(15)}{31 + 31 - 2}}} = 1.84$$

$H_0: \mu_A = \mu_B$ | $H_A: \mu_A > \mu_B$ | $\alpha = 0.05$

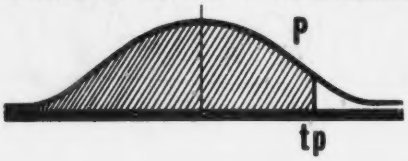
STATISTIC = 1.84 | t - DISTB. | df = 60



C.V. = $t_{.95}$(df=60) = 1.671

DEC. REJECT H_0

Percentiles of the t Distribution

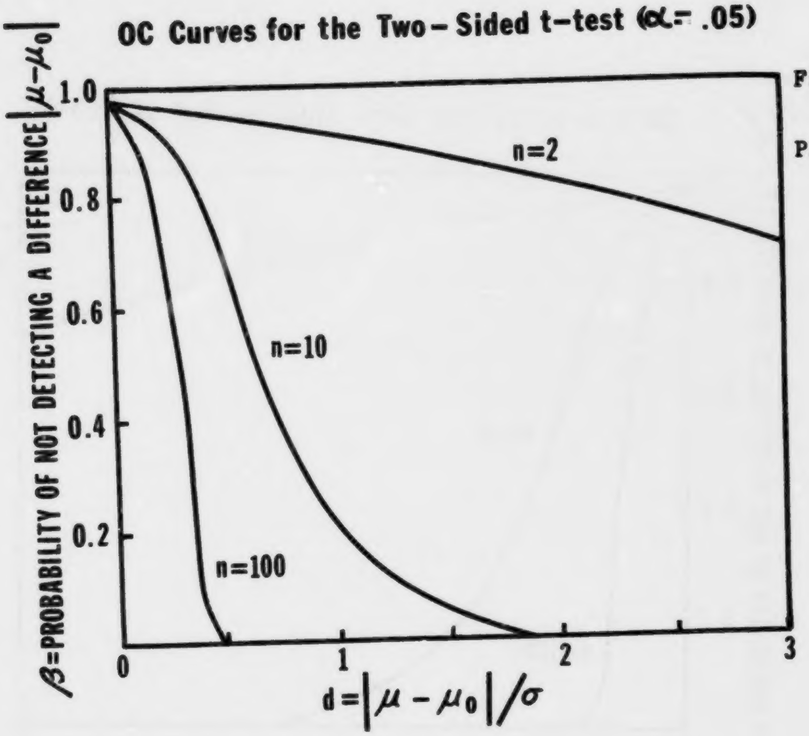


df	...	$t_{.95}$	$t_{.975}$	...
...				
24		1.711		
26			2.056	
...				
60		1.671		

Table A-4
T-5

GIVEN $n=10$ | $\mu-\mu_0=1$ | $\sigma=1$ | $\therefore d=1$

$\therefore \beta=0.2$ $1-\beta=0.8=$ POWER OF TEST = PROB. OF ACCEPTING H_A WHEN TRUE



OC Curves for the Two-Sided t-test ($\alpha = .05$)

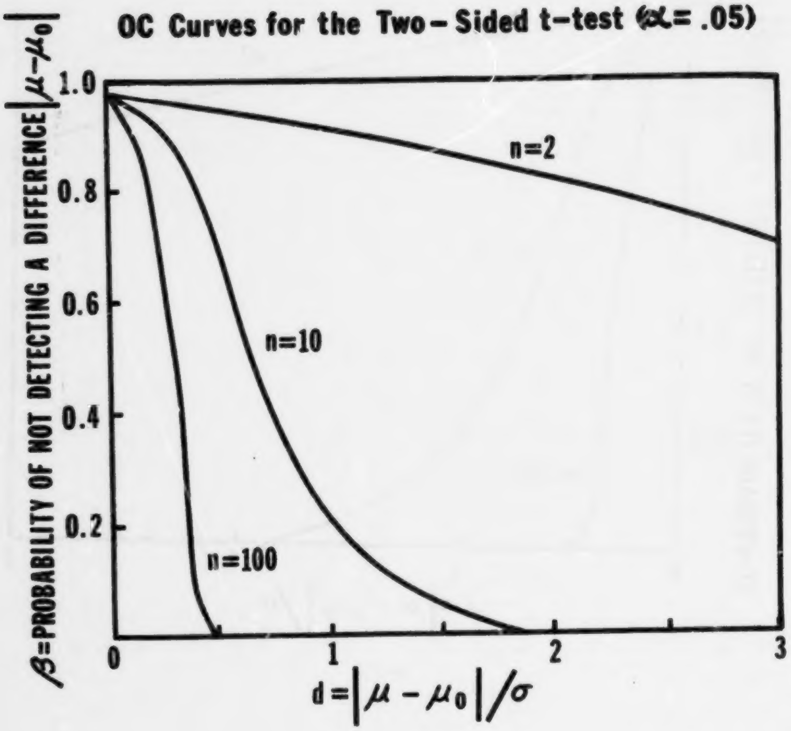
Figure 3-1

Page 3-6

GIVEN $n=10$ $\mu-\mu_0=1.5$ $\sigma=1$ $\therefore d=1.5$

$\beta = 0.03$ $\therefore 1-\beta = 0.97$

LARGER DIFF. ARE EASIER TO DETECT



OC Curves for the Two-Sided t-test ($\alpha = .05$)

Figure 3-1

Page 3-6

GIVEN: $n=10$: $\mu-\mu_0=0.5$: $\sigma=1$ ∴ $d=0.5$

$\beta=0.7$ ∴ $1-\beta=0.3$

PROB. OF DETECTING

SMALLER DIFFERENCE IS LESS

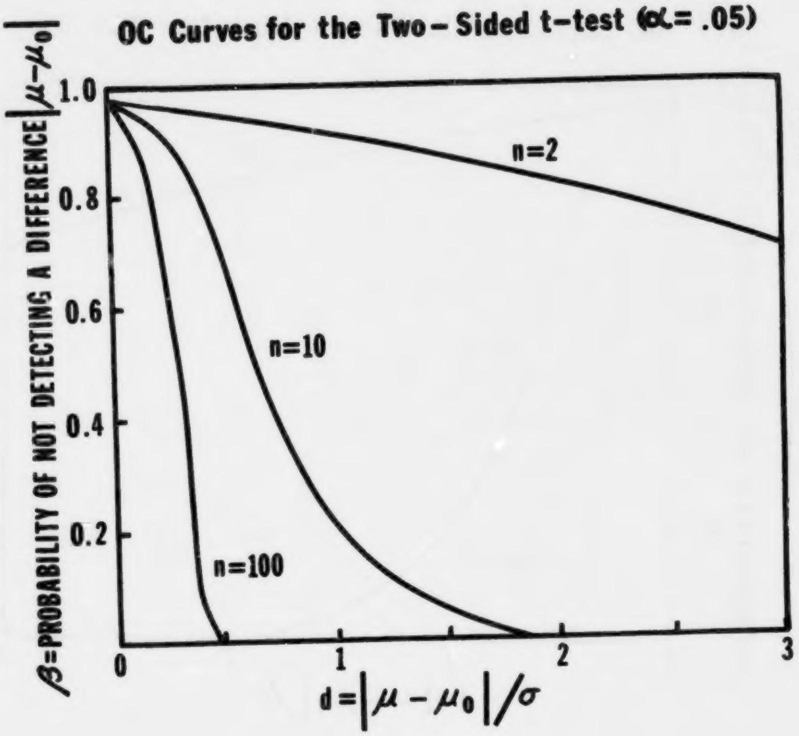

OC Curves for the Two-Sided t-test ($\alpha = .05$)

Figure 3-1
Page 3-6

GIVEN: $m=100$ | $\mu-\mu_0=0.5$ | $\sigma=1$ | $d=0.5$

β=ALMOST ZERO | $1-\beta$ = ALMOST ONE

INCREASING SAMPLE SIZE.

INCREASES PROB. OF DETECTION.

OC Curves for the Two-Sided t-test (α = .05)

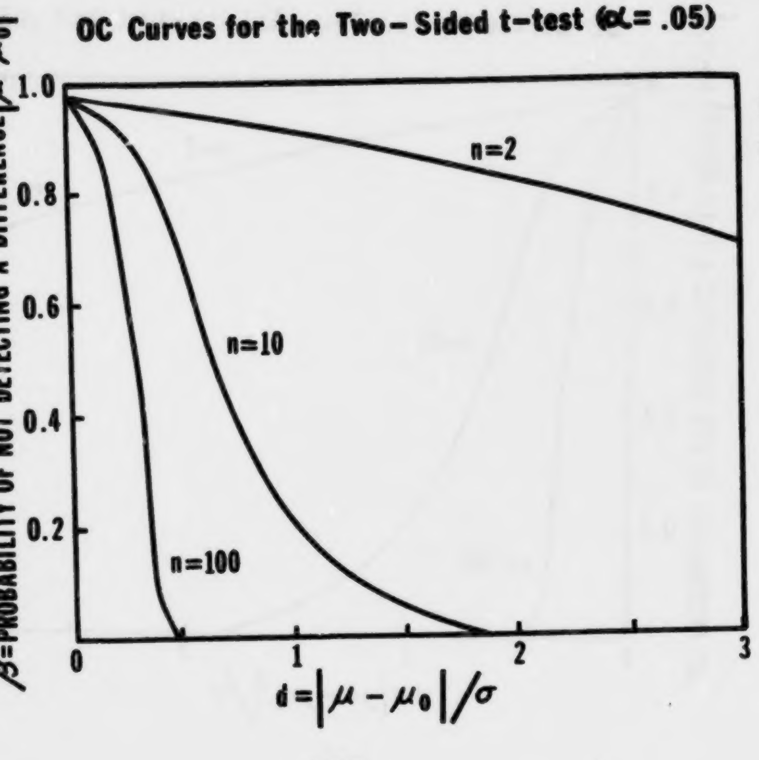


Figure 3-1

Page 3-6

GIVEN ¦ d (EST.) ¦ β ¦ FIND ¦ n

GIVEN ¦ d (COR.) ¦ n ¦ FIND ¦ β

see Table 3-1 Page 3-4

see Table 3-2 Page 3-22

see Article 3-3.1.4 paired observations two-sided test

see Article 3-3.2.4 paired observations one-sided test

see Article 3-4 for the k-sample problem

LET N_1=RAN. SAMP. NOR$_1$

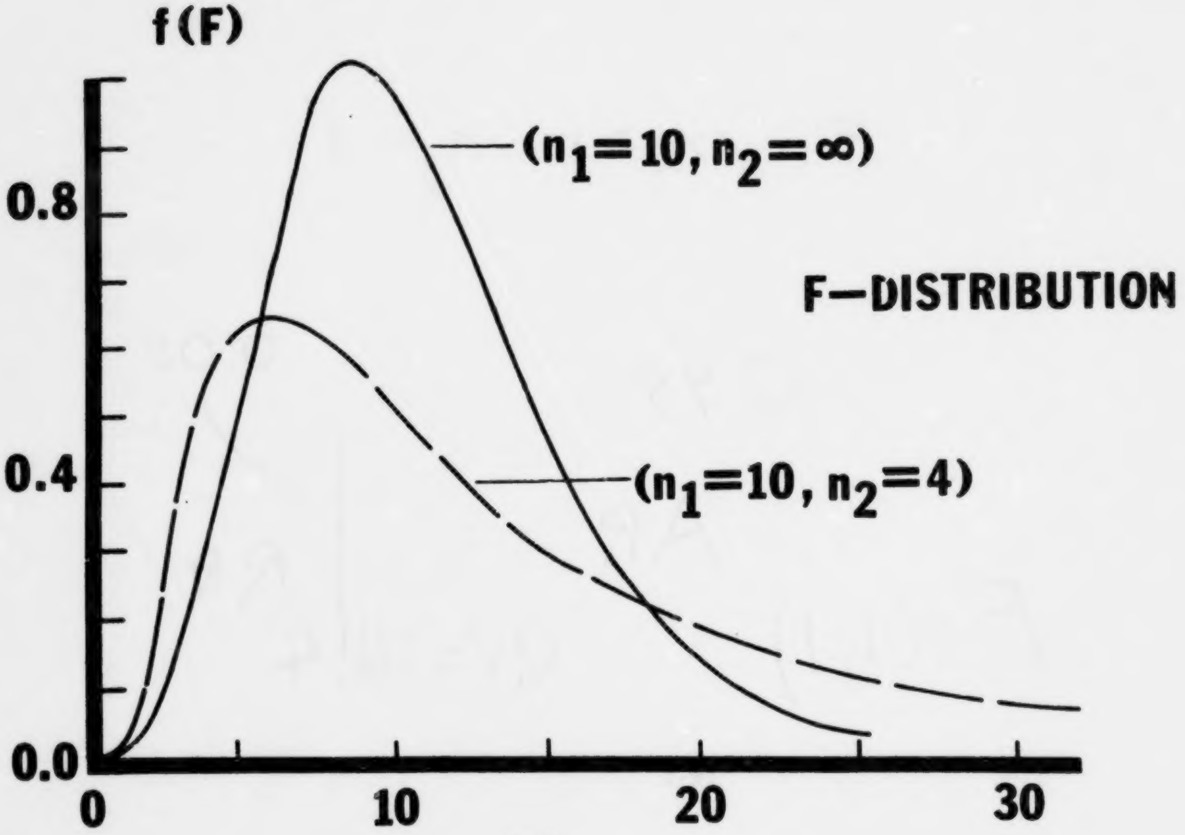
" N_2= " " NOR$_2$

EITHER NOR$_1$=NOR$_2$ | EQ. σ^2

$RV = \frac{s_1^2}{s_2^2}$ $F(n_1, n_2)$

$n_1 = N_1 - 1$ = df FOR NUM.

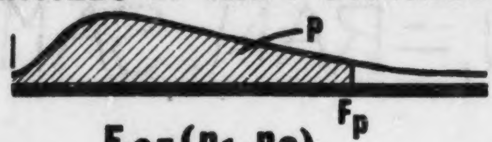
$n_2 = N_2 - 1$ = df FOR DEM.



1E-1

PERCENTILES OF THE F DISTRIBUTION

Table A-5
T-7



$F_{.95}(n_1,n_2)$

n_2 \ n_1	1	...	30	...	∞
1	161.4		250.1		254.3
2	18.51		19.46		19.50
30	4.17		1.84		1.62
⋮					
∞	3.84		1.46		1.00

F $n_1=1$ $n_2=1$

0.95

0.05

AR

RR.

F (1,1)

C.V=161.4

F $n_1=1$ $n_2=30$

0.95

0.05

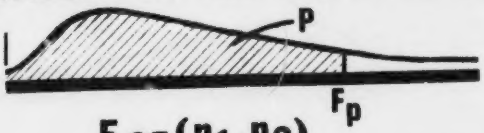
AR

RR.

$F_{0.95}(1,30) \longrightarrow C.V=4.17$

PERCENTILES OF THE F DISTRIBUTION

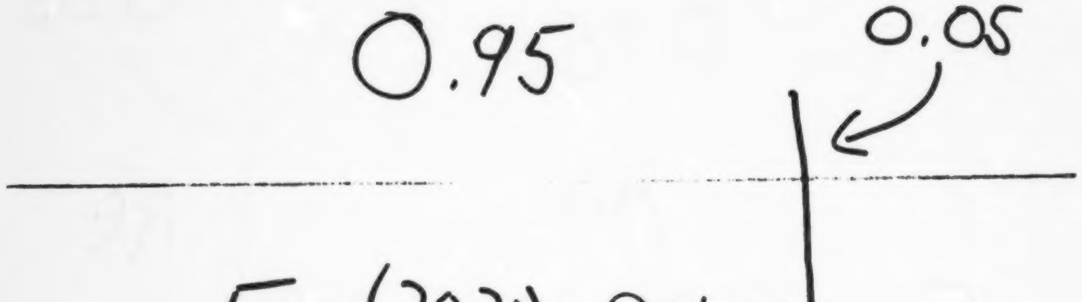
Table 4-5
T-7



$F_{.95}(n_1,n_2)$

n_2 \ n_1	1	. . .	30	. . .	∞
1	161.4		250.1		254.3
2	18.51		19.46		19.50
30	4.17		1.84		1.62
⋮					
∞	3.84		1.46		1.00

$F_{0.95}(30,30)$



$F_{0.95}(30,30) = C.V = 1.84$

PERCENTILES OF THE F DISTRIBUTION

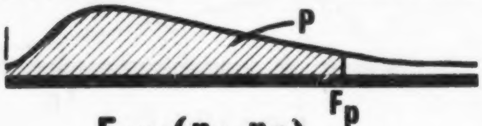


Table A-5
T-7

$F_{.95}(n_1, n_2)$

n_2 \ n_1	1	...	30	...	∞
1	161.4		250.1		254.3
2	18.51		19.46		19.50
30	4.17		1.84		1.62
⋮					
∞	3.84		1.46		1.00

One Sample Problem

	1	2	3
H_0	$\sigma = C$	$\sigma = C$	$\sigma = C$
H_A	$\sigma \neq C$	$\sigma < C$	$\sigma > C$

C = CONSTANT

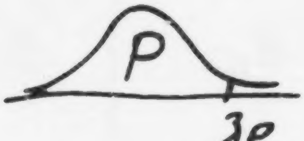
$H_0: \sigma = 10 \ (\sigma \leq 10)$, $H_A: \sigma > 10$, $\alpha = 0.05$

$\beta = 0.1$,

$\alpha = 0.05, 1-\alpha = 0.95, \beta = 0.1, 1-\beta = 0.9$

DEFINE $\lambda = \frac{\sigma}{c} = \frac{\sigma}{10} = 1.58$

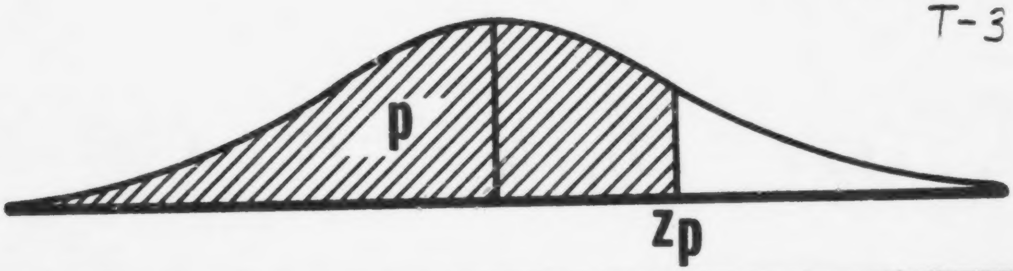
$$n = 1 + \frac{1}{2}\left(\frac{z_{1-\alpha} + \lambda z_{1-\beta}}{\lambda - 1}\right)^2$$



$$= 1 + \frac{1}{2}\left(\frac{1.645 + 1.58(1.282)}{1.58 - 1}\right)^2 = 21.025$$

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
z_p	1.282	1.645	1.960	3.090

$H_0: \sigma = 10 (\sigma \leq 10)$, $H_A: \sigma > 10$, $\alpha = 0.05$

$\beta = 0.1$, $n = 21$, $df = 20$, $s = 12.25$

$A_{.05} = 0.7979$

$s_L = A_{.05}(s) = 0.7979(12.25) = 9.77$

95% CONF. INT. FOR PAR. POP. σ IS

9.77
10
∞

DEC. DO NOT REJ. H_0

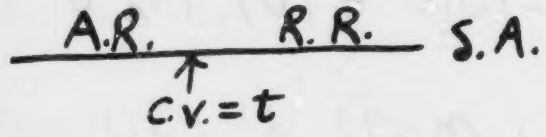
FACTORS FOR COMPUTING ONE-SIDED CONFIDENCE LIMITS FOR σ

Table A-21
T-30

DEG. OF FREEDOM	$A_{.05}$	$A_{.95}$
1	0.5103	15.947
2	0.5778	4.415
20	0.7979	1.358
⋮		
100	0.8968	1.133

ALSO — $A_{.025}$ $A_{.01}$ $A_{.005}$ $A_{.975}$ $A_{.99}$ $A_{.995}$

t-TEST μ



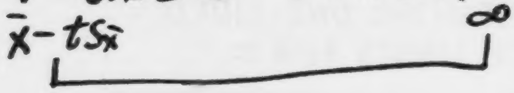
$H_0: \mu = \mu_0 \quad H_A: \mu > \mu_0$

IF $(\bar{x} - \mu_0)/s_{\bar{x}} < t$ DO NOT REJ. H_0

IF $(\bar{x} - \mu_0) < t\, s_{\bar{x}}$ " " " "

$\bar{x} - t s_{\bar{x}} < \mu_0$ " " " "

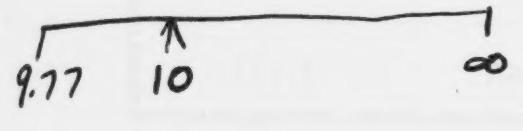
1-SIDED LOWER CONF. INT. FOR μ


$\bar{x} - t s_{\bar{x}}$
∞

IF $\bar{x} - t s_{\bar{x}} < \mu_0$ DO NOT REJ. H_0

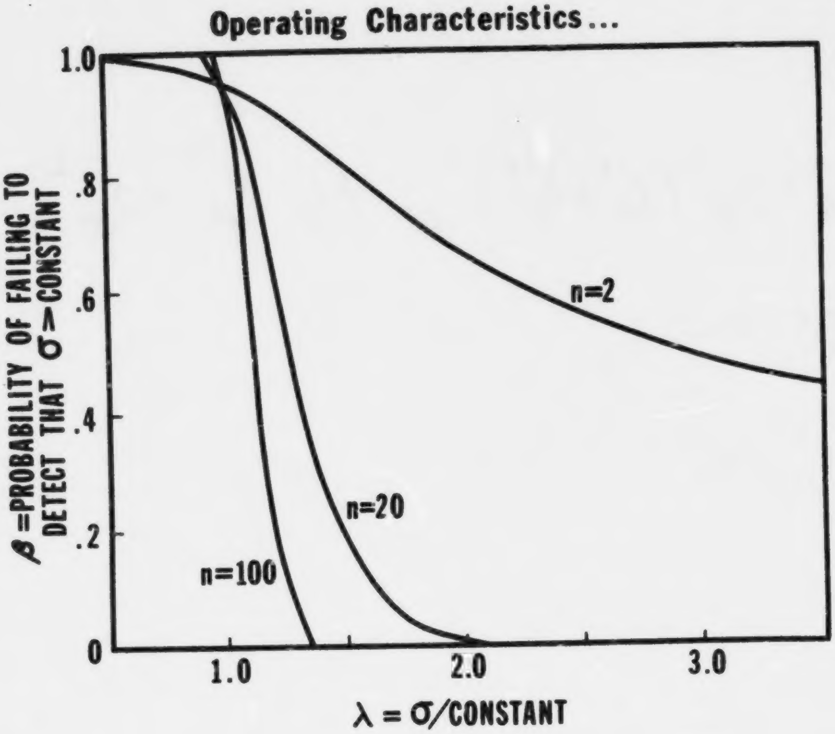
IF μ_0 INSIDE CONF. INT. DO NOT REJ. H_0

" " OUTSIDE " " REJ. H_0


9.77
10
∞

INSIDE $\therefore$ DO NOT REJ. H_0

Fig. 4-1
Page 4-4


Operating Characteristics...

Two Sample Problem

1	2
$H_0 : \sigma_A^2 = \sigma_B^2$ $H_A : \sigma_A^2 \neq \sigma_B^2$	$H_0 : \sigma_A^2 = \sigma_B^2$ $H_A : \sigma_A^2 > \sigma_B^2$

$H_0 : \sigma_A^2 = \sigma_B^2 \; (\sigma_A^2 \leq \sigma_B^2)$, $H_A : \sigma_A^2 > \sigma_B^2$

$\alpha = 0.05, \; \beta = 0.1,$

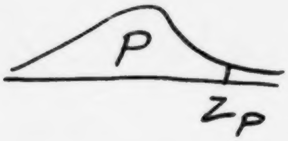
$F = s_A^2 / s_B^2 =$

$F_{1-\alpha}$ DIST. $(n_A - 1, n_B - 1)$,

$\alpha = 0.05 \quad \beta = 0.1$

DEFINE $\lambda = \sigma_A / \sigma_B = 1.73$

$$n = 2 + \left(\frac{z_{1-\alpha} + z_{1-\beta}}{\ln \lambda} \right)^2$$

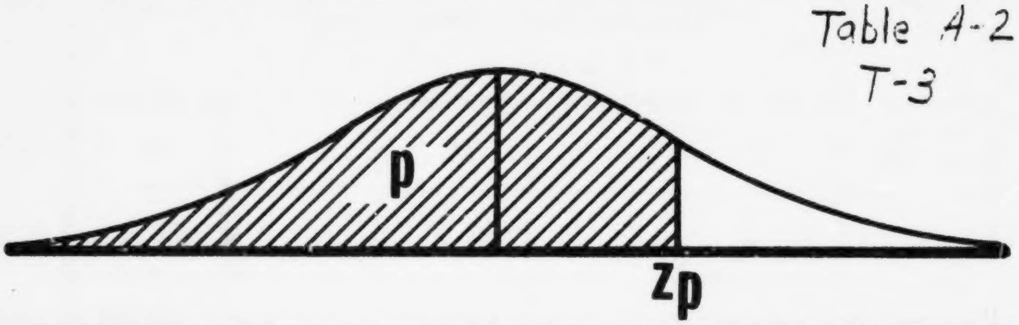


$$= 2 + \left(\frac{1.645 + 1.282}{\ln 1.73} \right)^2 = 2 + \left(\frac{2.927}{0.5481} \right)^2$$

$n = 30.5 \qquad n_A = n_B = 31$

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



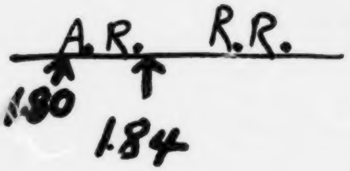
P	.9	.95	.975	.999
z_p	1.282	1.645	1.960	3.090

$H_0: \sigma_A^2 = \sigma_B^2 \ (\sigma_A^2 \le \sigma_B^2)$, $H_A: \sigma_A^2 > \sigma_B^2$

$\alpha = 0.05, \beta = 0.1, n_A = n_B = 31, s_A^2 = 9.5, s_B^2 = 5.28$

$F = s_A^2 / s_B^2 = 9.5/5.28 = 1.80$

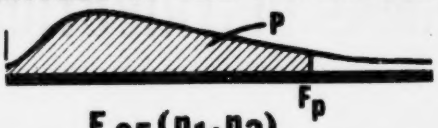
$F_{1-\alpha}$ DIST. $(n_A - 1, n_B - 1)$, $F_{.95}(30,30) = 1.84$



S.A., DE. DO NOT REJ. H_0

PERCENTILES OF THE F DISTRIBUTION

Table A-5
T-7

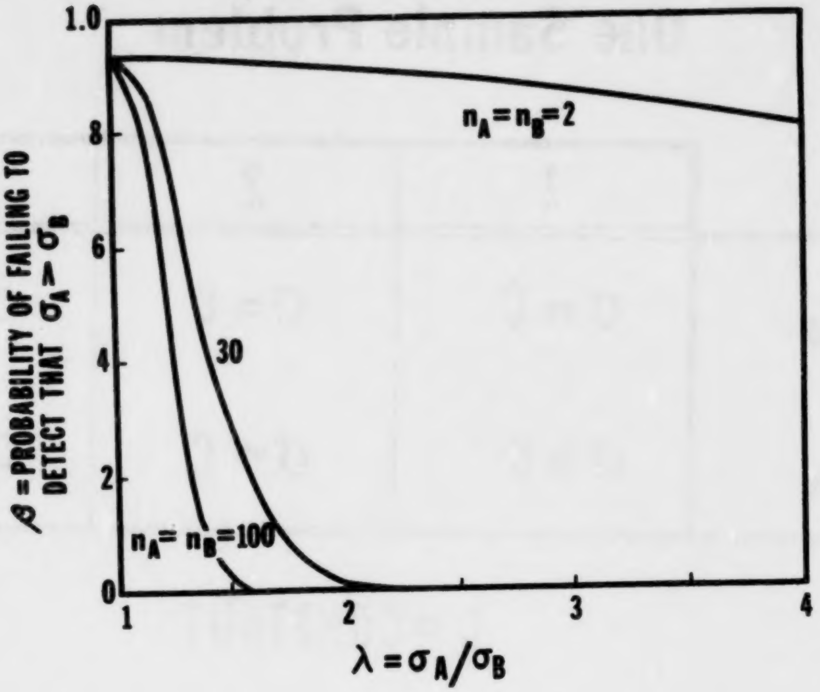


$F_{.95}(n_1, n_2)$

n_2 \ n_1	1	...	30	...	∞
1	161.4		250.1		254.3
2	18.51		19.46		19.50
30	4.17		1.84		1.62
⋮					
∞	3.84		1.46		1.00

Fig 1-3
Page 1-[illegible]

Operating Characteristics...



$H_0: \sigma_A^2 = \sigma_B^2 \ (\sigma_A^2 \leq \sigma_B^2)$, $H_A: \sigma_A^2 > \sigma_B^2$

$\alpha = 0.05$, $\beta = 0.1$, $n_A = n_B = 31$, $s_A^2 = 9.5$, $s_B^2 = 5.28$

$F = s_A^2 / s_B^2 = 9.5 / 5.28 = 1.80$

$F_{1-\alpha}$ DIST$(n_A - 1, n_B - 1)$, $F_{.95}(30, 30) = 1.84$

$\underset{\uparrow \; 1.80 \qquad \uparrow \; 1.84}{\underline{\text{A.R.} \qquad \text{R.R.}}}$ S.A., DE. DO NOT REJ. H_0

One Sample Problem

	1	2	3
H_0	$\sigma = C$	$\sigma = C$	$\sigma = C$
H_A	$\sigma \neq C$	$\sigma < C$	$\sigma > C$

C = CONSTANT

FACTORS FOR COMPUTING ONE-SIDED CONFIDENCE LIMITS FOR σ

Table A-21
T-36

DEG. OF FREEDOM	$A_{.05}$	$A_{.95}$
1	0.5103	15.947
2	0.5778	4.415
20	0.7979	1.358
⋮		
100	0.8968	1.133

ALSO — $A_{.025}$ $A_{.01}$ $A_{.005}$ $A_{.975}$ $A_{.99}$ $A_{.995}$

C. I. 1-S. A_L ∞

One Sample Problem

	1	2	3
H_0	$\sigma = C$	$\sigma = C$	$\sigma = C$
H_A	$\sigma \neq C$	$\sigma < C$	$\sigma > C$

C = CONSTANT

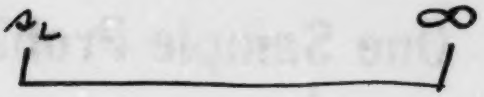
FACTORS FOR COMPUTING ONE-SIDED CONFIDENCE LIMITS FOR σ

Table A-21
T-36

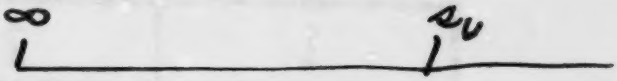
DEG. OF FREEDOM	$A_{.05}$	$A_{.95}$
1	0.5103	15.947
2	0.5778	4.415
20	0.7979	1.358
⋮		
100	0.8968	1.133

ALSO — $A_{.025}$ $A_{.01}$ $A_{.005}$ $A_{.975}$ $A_{.99}$ $A_{.995}$

C. I. 1-S.



C. I. 1-S.



One Sample Problem

	1	2	3
H_0	$\sigma = C$	$\sigma = C$	$\sigma = C$
H_A	$\sigma \neq C$	$\sigma < C$	$\sigma > C$

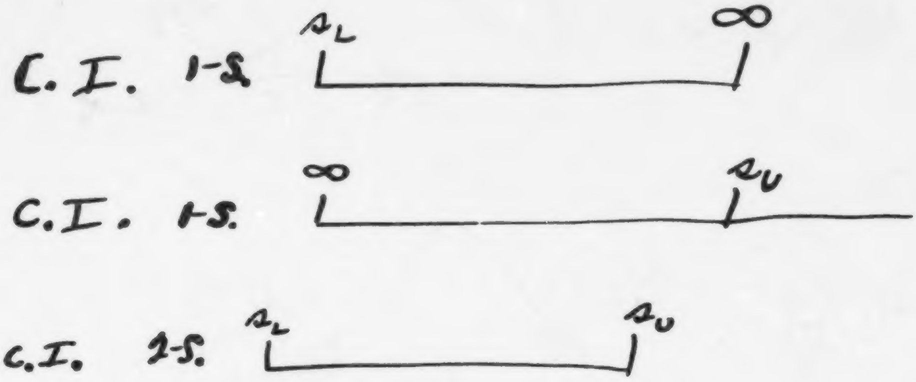
C = CONSTANT

FACTORS FOR COMPUTING TWO-SIDED CONFIDENCE LIMITS FOR σ

Table A-20
T-34

DEG. OF FREEDOM ν	$\alpha = 0.05$	
	B_U	B_L
1	17.79	0.3576
2	4.859	0.4581
3	3.183	0.5178
⋮		
100	1.157	0.8757

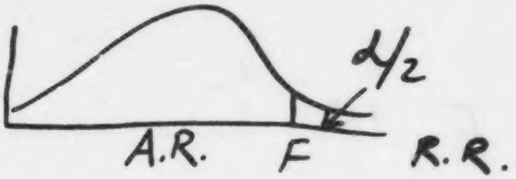
ALSO $\alpha = 0.01, 0.001$


C.I. 1-S.
s_L
∞
C.I. 1-S.
∞
s_U
C.I. 2-S.
s_L
s_U

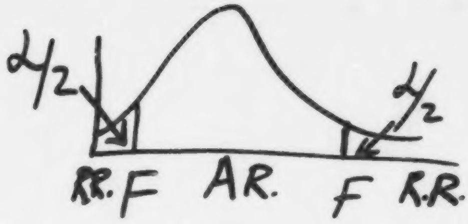
Two Sample Problem

1	2
$H_0: \sigma_A^2 = \sigma_B^2$ $H_A: \sigma_A^2 \neq \sigma_B^2$	$H_0: \sigma_A^2 = \sigma_B^2$ $H_A: \sigma_A^2 > \sigma_B^2$

$$F = \frac{LAR.\ V}{SM.\ V} \geq 1$$

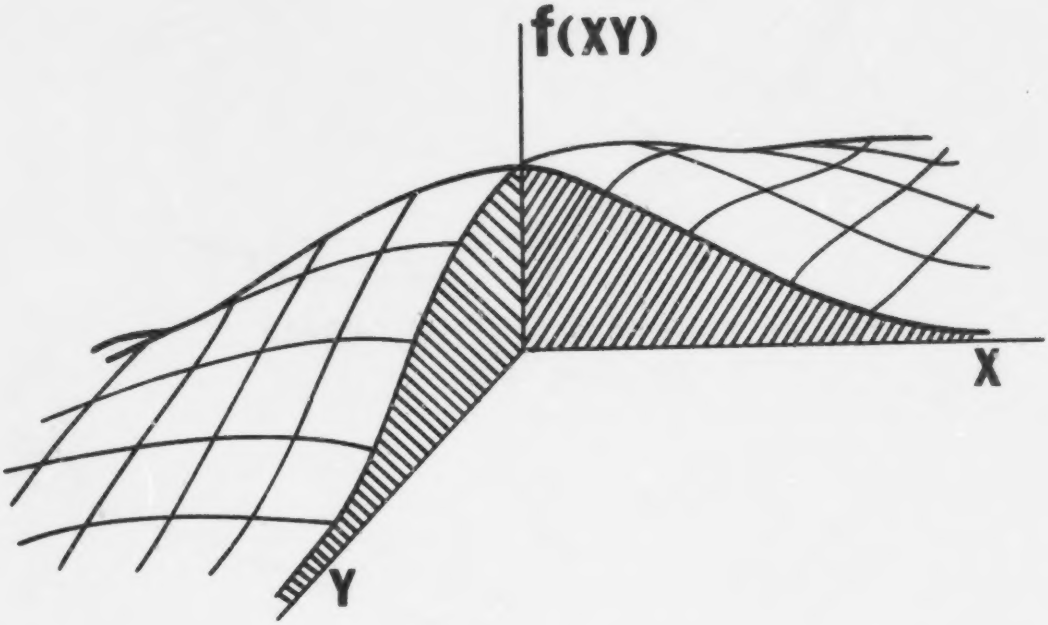


$$F = \frac{\sigma_A^2}{\sigma_B^2}$$

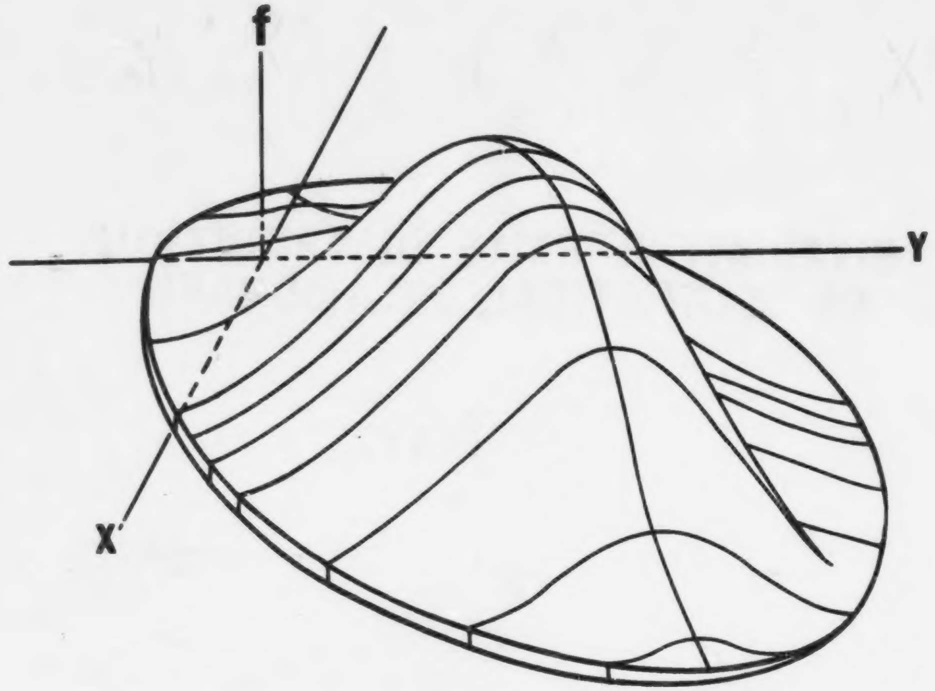

α/2
α/2
R.R.
F
A.R.
F
R.R.

$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$

BIVARIATE NORMAL DISTRIBUTION OR STATISTICAL RELATIONSHIP

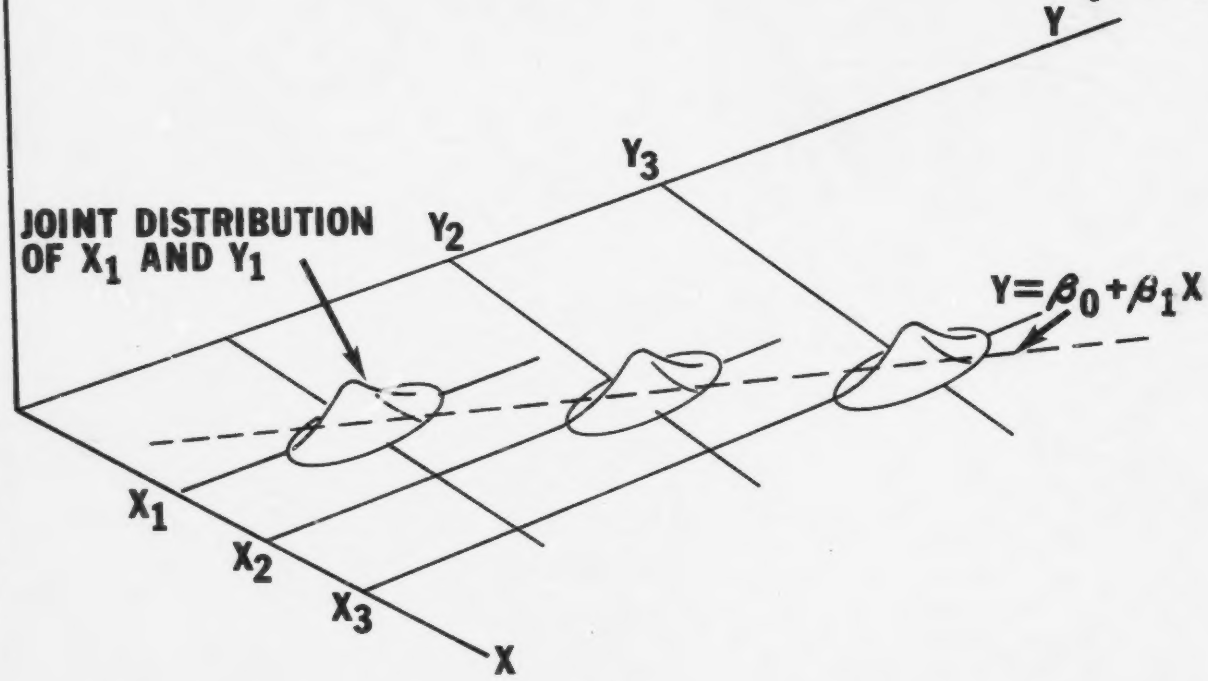


TWO-DIMENSIONAL NORMAL DISTRIBUTION
OR STATISTICAL RELATIONSHIP

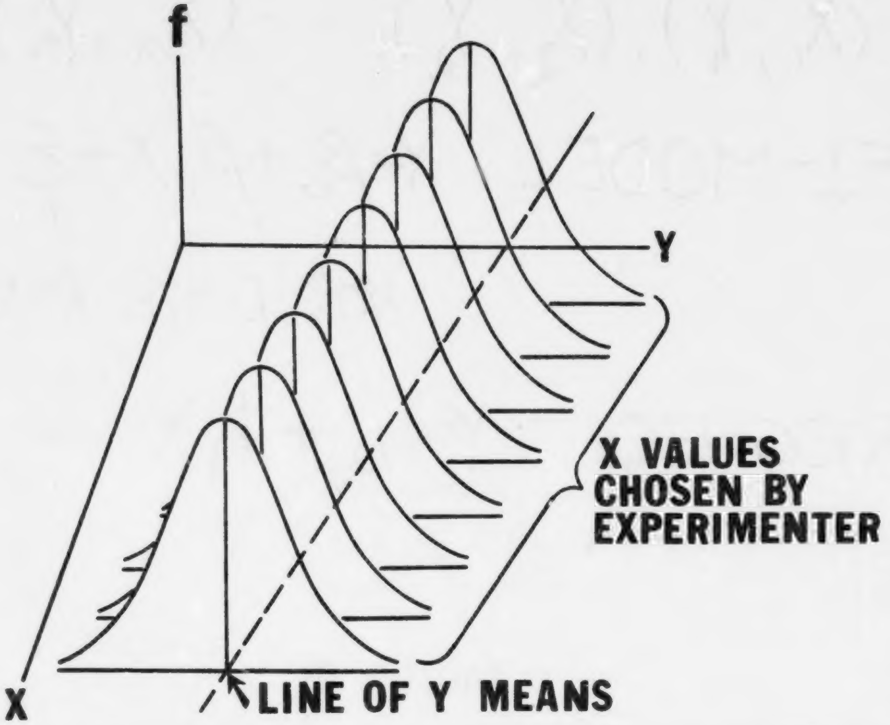


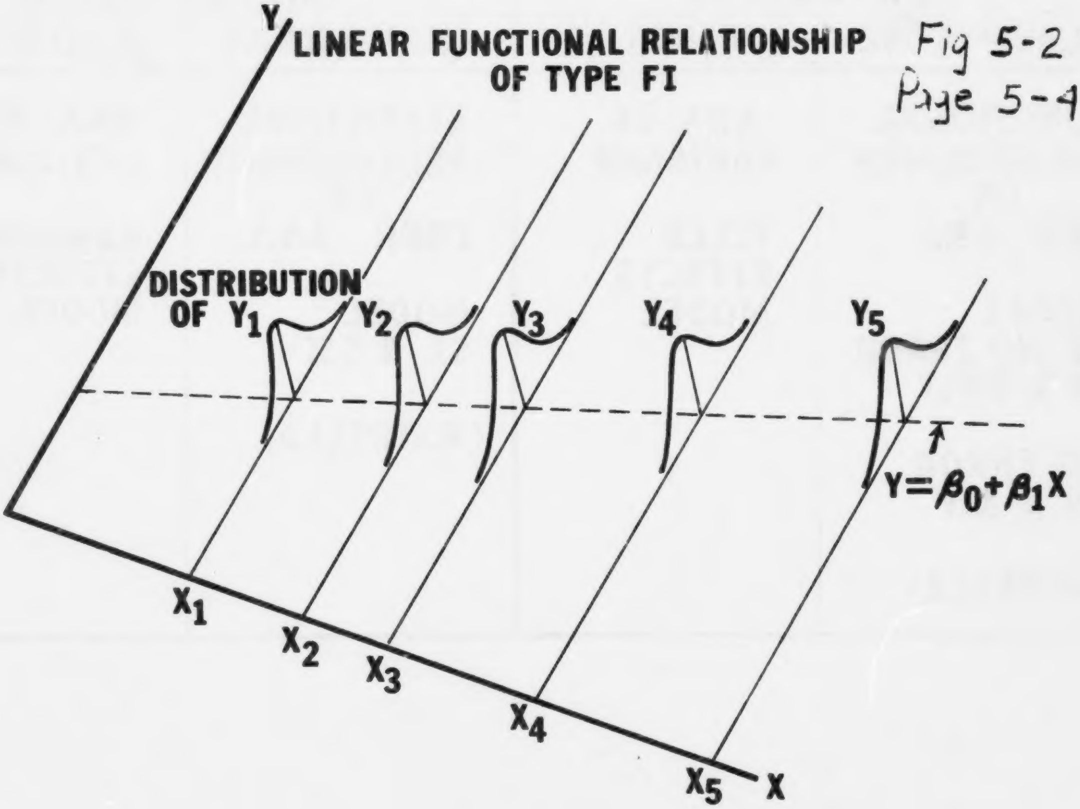
LINEAR FUNCTIONAL RELATIONSHIP
OF TYPE FII

Fig. 5-3
Page 5-5


Y
Y_3
Y_2
JOINT DISTRIBUTION OF X_1 AND Y_1
$Y=\beta_0+\beta_1 X$
X_1
X_2
X_3
X

FUNCTIONAL RELATIONSHIP FI


f
Y
X VALUES
CHOSEN BY
EXPERIMENTER
X
LINE OF Y MEANS


Y
DISTRIBUTION
OF Y_1
Y_2
Y_3
Y_4
Y_5
$Y=\beta_0+\beta_1X$
X_1
X_2
X_3
X_4
X_5
X

LINEAR FUNCTIONAL RELATIONSHIP
OF TYPE FI

Fig 5-2
Page 5-4

$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$

$$\text{FI-MODEL} \quad Y = \beta_0 + \beta_1 X + \underset{\substack{\uparrow \\ \text{NOR. DISTB. RAN. ERR.}}}{\varepsilon}$$

$$\text{REG. EQ.} \quad Y = b_0 + b_1 X$$

X FACTOR IS

CONTROLLED		UNCONTROLLED	
QUANTITATIVE	QUALITATIVE	QUANTITATIVE	QUALITATIVE
FUNCTIONAL RELATIONSHIP OR REG. ANA. MODEL FI NO ERROR IN X ONLY FII ERROR IN X & Y (NATRELLA)	ANA. OF VARIANCE FIXED EFFECTS MODEL	STATISTICAL RELATIONSHIP OR CORR. ANA. MODEL SI & SII (NATRELLA)	ANA. OF VARIANCE RANDOM EFFECTS MODEL

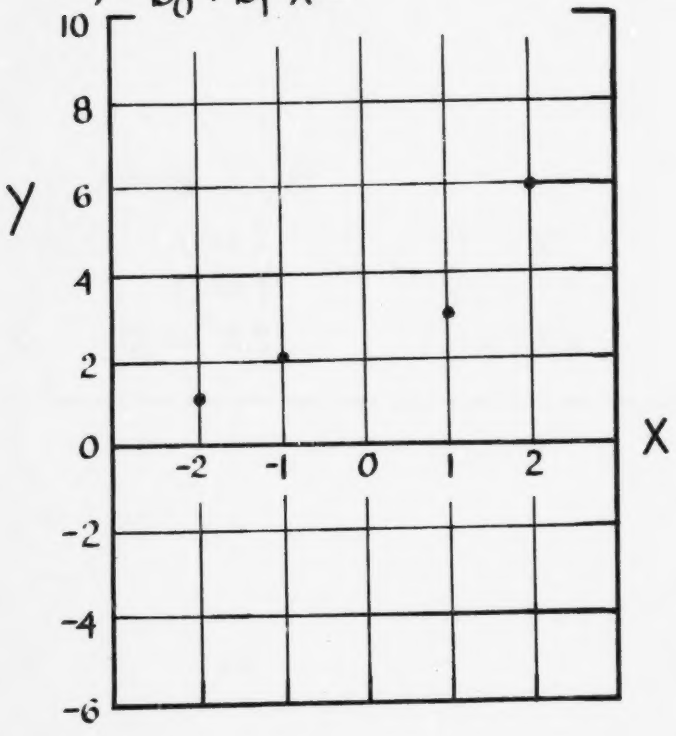
MODEL FI WKSHEET, 5-13

X	Y
−2	1
−1	2
1	3
2	6

$Y = \beta_0 + \beta_1 X + \underset{\text{ERROR TERM}}{\varepsilon}$ (MODEL)

MODEL FI

$Y = b_0 + b_1 X =$



MODEL FI WKSHEET, 5-13

X^2	X	XY	Y	Y^2
4	-2	-2	1	1
1	-1	-2	2	4
1	1	3	3	9
4	2	12	6	36
10	$\Sigma X = 0$	11	$\Sigma Y = 12$	50
ΣX^2	$\bar{X} = 0$	ΣXY	$\bar{Y} = 3$	ΣY^2

DATA BANK

$\Sigma X = 0$	$\bar{X} = 0$	$n = 4$
$\Sigma Y = 12$	$\bar{Y} = 3$	
$\Sigma XY = 11$	$\Sigma X^2 = 10$	$\Sigma Y^2 = 50$

$\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

$- (\Sigma X)^2/n = 0$ $- (\Sigma Y)^2/n = 36$

$S_{xx} = 10$ $S_{yy} = 14$

$\Sigma XY = 11$ $b_1 = \frac{S_{xy}}{S_{xx}} = \frac{11}{10} = 1.1$

$- (\Sigma X)(\Sigma Y)/n = 0$ $b_0 = \bar{Y} - b_1\bar{X} = 3 - 0 = 3$

$S_{xy} = 11$

DATA BANK

$\Sigma X = 0$	$\bar{X} = 0$	$n = 4$
$\Sigma Y = 12$	$\bar{Y} = 3$	
$\Sigma XY = 11$	$\Sigma X^2 = 10$	$\Sigma Y^2 = 50$

$S_{xx} = 10$		$S_{yy} = 14$
$S_{xy} = 11$	$b_1 = 1.1$	$b_0 = 3$

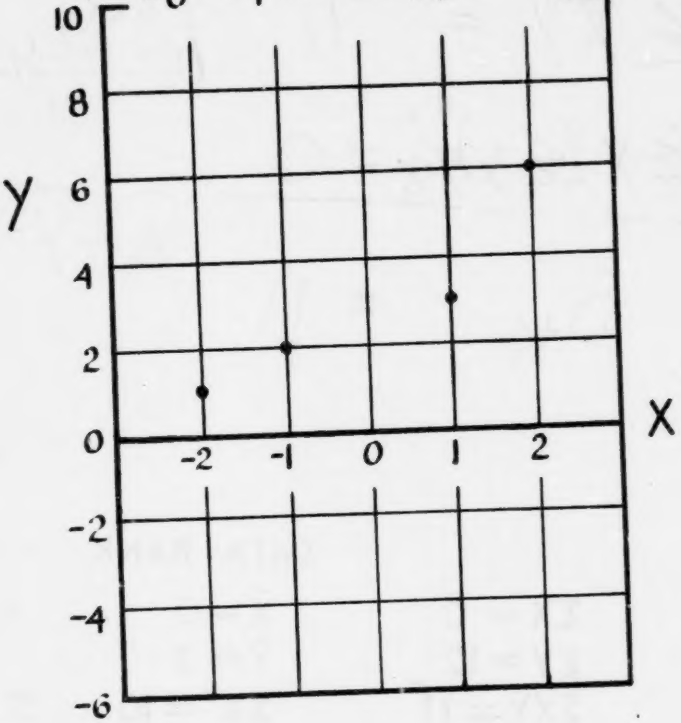
$S_{xx} = 10$ $\qquad$ $S_{yy} = 14$

$S_{xy} = 11$ $\qquad$ $b_1 = 1.1$ $\qquad$ $b_0 = 3$

$Y = \beta_0 + \beta_1 X + \varepsilon$ (MODEL) ← ERROR TERM

MODEL FI

$Y = b_0 + b_1 X = 3 + 1.1X$



$Y = 3 + 1.1X$ | $X = -3$ | $Y = -0.3$ | $(-3, -0.3)$

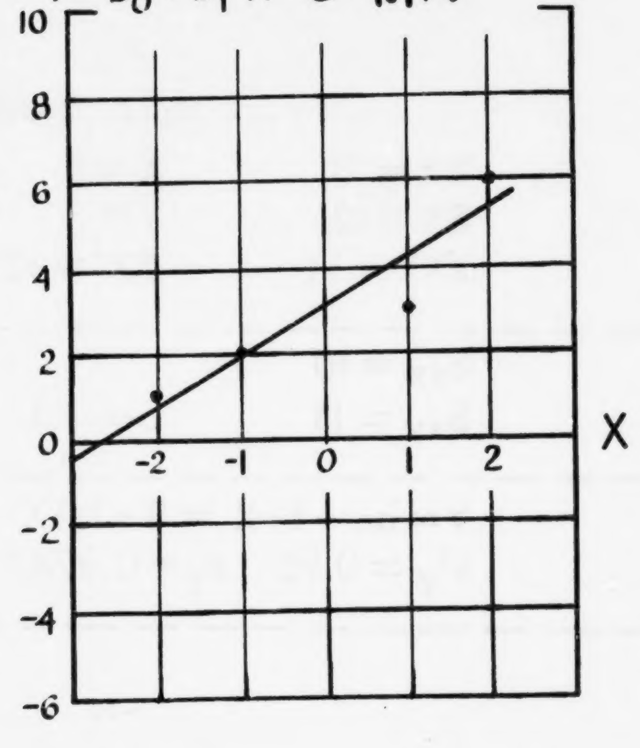
$X = 3$ | $Y = 6.3$ | $(3, 6.3)$ | $(0, 3)$

MODEL FI

$Y = \beta_0 + \beta_1 X + \varepsilon$ (MODEL)

ERROR TERM

$Y = b_0 + b_1 X = 3 + 1.1X$



$Y = 3 + 1.1X \quad X = -3 \quad Y = -0.3 \quad (-3, -0.3)$

$X = 3 \quad Y = 6.3 \quad (3, 6.3) \quad (0, 3)$

$$s^2_Y = \frac{1}{n-2}\left[S_{yy} - \frac{S^2_{xy}}{S_{xx}}\right] = \frac{1}{2}\left[14 - \frac{11^2}{10}\right] = 0.95$$

$$s_Y = \sqrt{0.95} = 0.9747$$

$$s^2_{b_1} = \frac{s^2_Y}{S_{xx}} = \frac{0.95}{10} = 0.095 \quad s_{b_1} = 0.3082$$

DATA BANK

$\Sigma X = 0 \qquad \bar{X} = 0 \qquad n = 4$

$\Sigma Y = 12 \qquad \bar{Y} = 3$

$\Sigma XY = 11 \qquad \Sigma X^2 = 10 \qquad \Sigma Y^2 = 50$

$S_{xx} = 10 \qquad S_{yy} = 14$

$S_{xy} = 11 \qquad b_1 = 1.1 \qquad b_0 = 3$

$Y = b_0 + b_1 X = 3 + 1.1X$

$s^2_Y = 0.95 \qquad s_Y = 0.9747 \qquad s_{b_1} = 0.3082$

$$s^2_{b_0} = s^2_Y\left[\frac{1}{n}+\frac{\bar{X}^2}{S_{xx}}\right] = 0.95\left[\frac{1}{4}+\frac{0^2}{S_{xx}}\right]$$

$s^2_{b_0} = 0.2375$ ⋮ $s_{b_0} = 0.4873$

DATA BANK

$\Sigma X = 0$ $\bar{X} = 0$ $n = 4$
$\Sigma Y = 12$ $\bar{Y} = 3$
$\Sigma XY = 11$ $\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

- - - - - - - - - - - - - - - - - - -

$S_{xx} = 10$ $S_{yy} = 14$
$S_{xy} = 11$ $b_1 = 1.1$ $b_0 = 3$

- - - - - - - - - - - - - - - - - - -

$Y = b_0 + b_1 X = 3 + 1.1X$
$s^2_Y = 0.95$ $s_Y = 0.9747$ $s_{b_1} = 0.3082$

- - - - - - - - - - - - - - - - - - -

$s_{b_0} = 0.4873$ $t_{.975}[df = 2] = 4.303$

PAGE T-5

Percentiles of the t Distribution

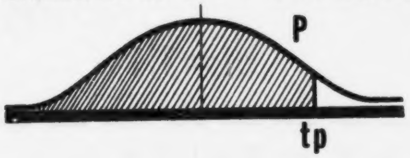


Table 4-4
T-5

df	...	$t_{.95}$	$t_{.975}$	...
...	.			
2			4.303	
...				

100(1−α)% CONF. INT. FOR $\bar{Y}$ AT

$X = X'$ | $\alpha = 0.05$ | LET $X' = 0$ | $\bar{Y}_{X'=0} = 3$

$$W_2 = t_{1-0.5\alpha}(df = n-2)\, s_y \sqrt{\frac{1}{n} + \frac{(X' - \bar{X})^2}{S_{xx}}}$$

$$= 4.3(0.975)\sqrt{\frac{1}{4} + \frac{(0-0)^2}{S_{xx}}} = 2.1$$

$$L: \bar{Y} \pm W_2 = 3 \pm 2.1 = 0.9$$
$$= 5.1$$

(0, 0.9)
(0, 5.1)

DATA BANK

$\Sigma X = 0$ $\bar{X} = 0$ $n = 4$
$\Sigma Y = 12$ $\bar{Y} = 3$
$\Sigma XY = 11$ $\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

$S_{xx} = 10$ $S_{yy} = 14$
$S_{xy} = 11$ $b_1 = 1.1$ $b_0 = 3$

$Y = b_0 + b_1 X = 3 + 1.1X$
$s^2_Y = 0.95$ $s_Y = 0.9747$ $s_{b_1} = 0.3082$

$s_{b_0} = 0.4873$ $t_{.975}[df = 2] = 4.303$

PAGE T-5

100(1-α)% CONF. INT. FOR $\bar{Y}$ AT

$X = X'$ | $\alpha = 0.05$ | LET $X' = 0$ | $\bar{Y}_{X'=0} = 3$

$$W_2 = t_{1-0.5\alpha}(df = n-2)\, s_y \sqrt{\frac{1}{n} + \frac{(x'-\bar{x})^2}{S_{xx}}}$$

$$= 4.3(0.975)\sqrt{\frac{1}{4} + \frac{(0-0)^2}{S_{xx}}} = 2.1$$

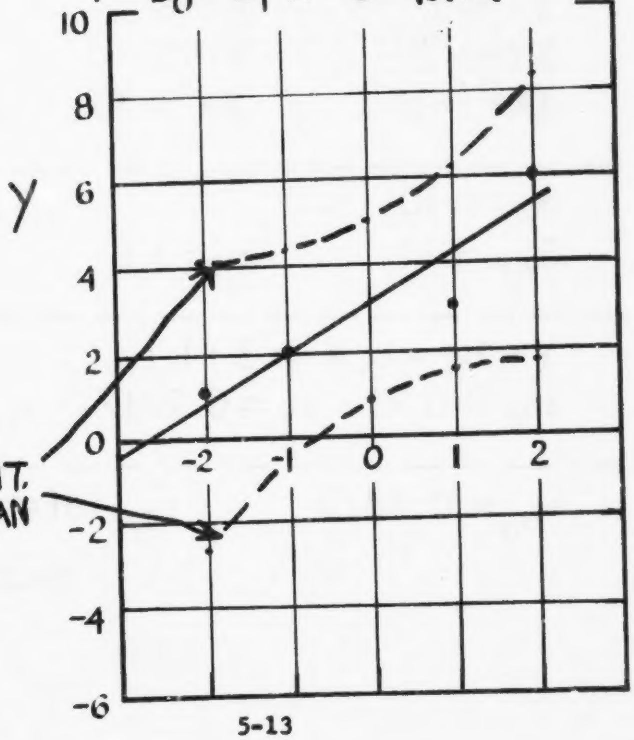
$L: \bar{Y} \pm W_2 = 3 \pm 2.1 = 0.9$ | $(0, 0.9)$

$= 5.1$ | $(0, 5.1)$

$Y = \beta_0 + \beta_1 X + \varepsilon$ (MODEL) ← ERROR TERM

$Y = b_0 + b_1 X = 3 + 1.1X$

MODEL FI



95% CONF. INT. FOR MEAN

100(1-α)% CONF. INT. AT $X = X'$ FOR

IND. Y's (TOL LIM.) | $\alpha = 0.05$ | $X' = 0$ | $\bar{Y}_{x'=0} = 3$

$$W_3 = t_{1-0.5\alpha}(df = n-2)\, s_y \sqrt{1 + \frac{1}{n} + \frac{(X' - \bar{X})^2}{S_{xx}}}$$

$$= 4.3(0.975)\sqrt{1 + \frac{1}{4} + \frac{(0-0)^2}{S_{xx}}} = 4.7$$

$L: \bar{Y} \pm W_3 = 3 \pm 4.7 = -1.7$ | $(0, -1.7)$

$= 7.7$ | $(0, 7.7)$

DATA BANK

$\Sigma X = 0$ $\bar{X} = 0$ $n = 4$

$\Sigma Y = 12$ $\bar{Y} = 3$

$\Sigma XY = 11$ $\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

$S_{XX} = 10$ $S_{YY} = 14$

$S_{XY} = 11$ $b_1 = 1.1$ $b_0 = 3$

$Y = b_0 + b_1 X = 3 + 1.1X$

$s^2_Y = 0.95$ $s_Y = 0.9747$ $s_{b_1} = 0.3082$

$s_{b_0} = 0.4873$ $t_{.975}[df = 2] = 4.303$

PAGE T-5

100(1−α)% CONF. INT. AT X= X′ FOR IND. Y′s (TOL LIM.) | $\alpha = 0.05$ | $X' = 0$ | $\bar{Y}_{X'=0} = 3$

$$W_3 = t_{1-.5\alpha}(df = n-2)\, s_y \sqrt{1 + \frac{1}{n} + \frac{(X' - \bar{X})^2}{S_{xx}}}$$

$$= 4.3(0.975)\sqrt{1 + \frac{1}{4} + \frac{(0-0)^2}{S_{xx}}} = 4.7$$

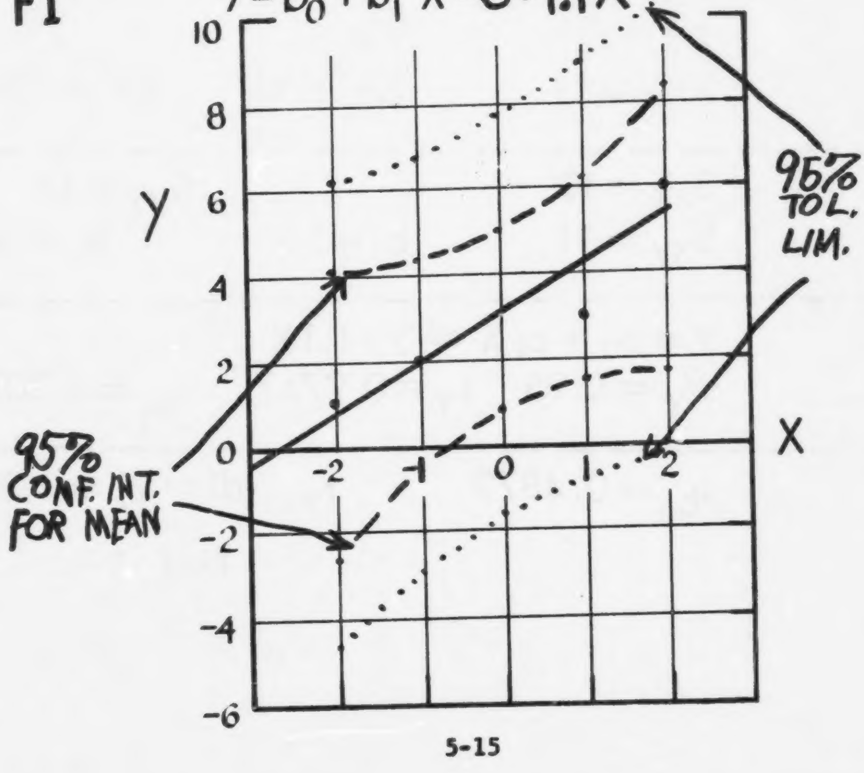
$$L: \bar{Y} \pm W_3 = 3 \pm 4.7 = -1.7 \qquad (0, -1.7)$$

$$= 7.7 \qquad (0, 7.7)$$

$Y = \beta_0 + \beta_1 X + \varepsilon$ (MODEL) ← ERROR TERM

$Y = b_0 + b_1 X = 3 + 1.1X$

MODEL FI

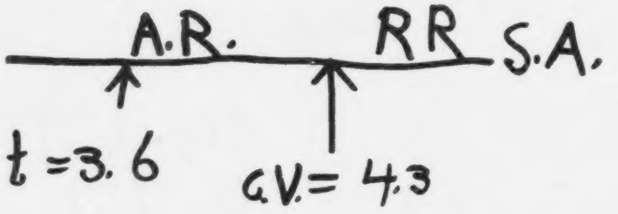


$H_0 : \beta_1 = \text{CONST.}$ ¦ STAT. $= (b_1 - \text{CONST})/s_{b_1}$

t-DIST. (df $= n-2$) ¦ $\alpha = 0.05$ ¦ $H_0 : \beta_1 = 0$

$H_A : \beta_1 \neq 0$ ¦ C. V $= t_{1-0.5\alpha}(2) = t_{0.975} = 4.3$

$t = |1.1 - 0| / 0.3082 = 3.6$



DO NOT

REJ. H_0

DATA BANK

$\Sigma X = 0$ $\bar{X} = 0$ $n = 4$
$\Sigma Y = 12$ $\bar{Y} = 3$
$\Sigma XY = 11$ $\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

$S_{xx} = 10$ $S_{yy} = 14$
$S_{xy} = 11$ $b_1 = 1.1$ $b_0 = 3$

$Y = b_0 + b_1 X = 3 + 1.1X$
$s^2_Y = 0.95$ $s_Y = 0.9747$ $s_{b_1} = 0.3082$

$s_{b_0} = 0.4873$ $t_{.975}[df = 2] = 4.303$

PAGE T-5

100(1-α)% CONF. INT. FOR β_1

$$b_1 - t_{1-0.5\alpha}(df = n-2) s_{b_1} < \beta_1 < b_1 + t s_{b_1}$$

$$-0.23 < \beta_1 < 2.43$$

$$\alpha = 0.05$$

$\beta_1 = 0$ IN INT. → DO NOT REJ. H_0

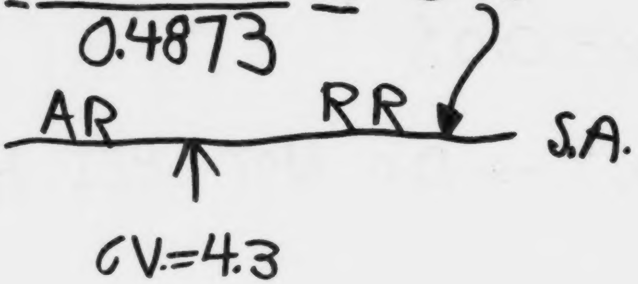
$H_0: \beta_0 = \text{CONST.}$ ┆ $\text{STAT} = (b_0 - \text{CONST})/s_{b_0}$

t- DISTB. $(df = n-2)$ ┆ $\alpha = 0.05$

$H_0: \beta_0 = 0$ ┆ $H_A: \beta_0 \neq 0$ ┆ $t_{0.975}(2) = 4.3$

$$t = \frac{|3-0|}{0.4873} = 6.16$$

REJ. H_0



$\beta_0 \neq 0$

DATA BANK

$\Sigma X = 0$ $\bar{X} = 0$ $n = 4$
$\Sigma Y = 12$ $\bar{Y} = 3$
$\Sigma XY = 11$ $\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

- -

$S_{xx} = 10$ $S_{yy} = 14$
$S_{xy} = 11$ $b_1 = 1.1$ $b_0 = 3$

- -

$Y = b_0 + b_1 X = 3 + 1.1X$
$s^2_Y = 0.95$ $s_Y = 0.9747$ $s_{b_1} = 0.3082$

- -

$s_{b_0} = 0.4873$ $t_{.975}[df = 2] = 4.303$

PAGE T-5

100(1-α)% CONF. INT. FOR β_0

$$b_0 - t_{1-0.5\alpha}(df=n-2)\, s_{b_0} < \beta_0 < b_0 + t\, s_{b_0}$$

$$0.9 < \beta_0 < 5.09$$

$\alpha = 0.05$ $\beta_0 = 0$ NOT IN

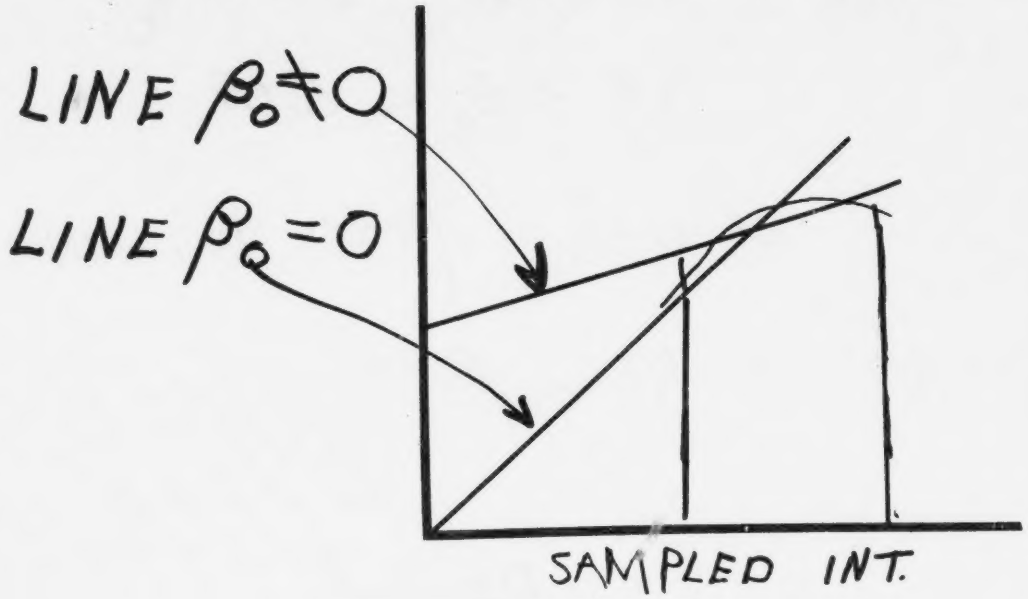
INT. → REJ. $\beta_0 = 0$

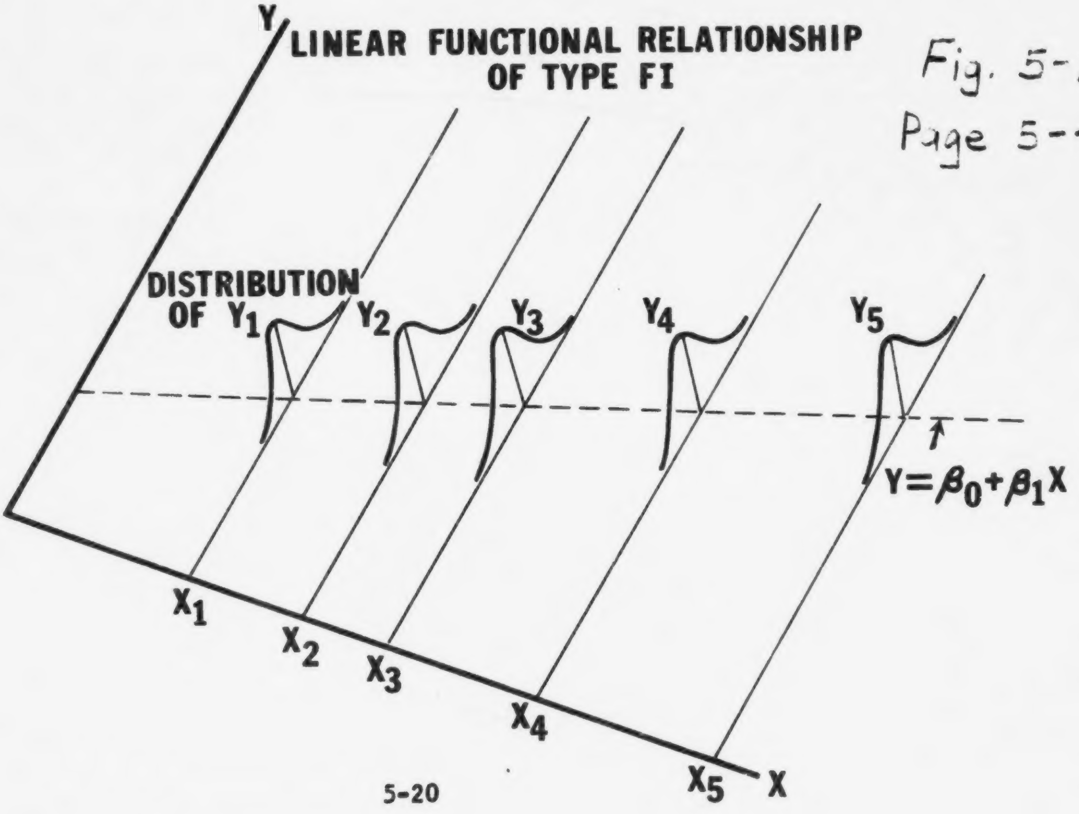
STUDENT

EXERCISE

MODEL: $Y=\beta_0+\beta_1X+\varepsilon$

IF KNOW: $\beta_0=0$ ⋮ SPEC. FORM.


LINE $\beta_0\neq0$
LINE $\beta_0=0$
SAMPLED INT.


LINEAR FUNCTIONAL RELATIONSHIP OF TYPE FI

Y
DISTRIBUTION OF Y_1
Y_2
Y_3
Y_4
Y_5
$Y=\beta_0+\beta_1X$
X_1
X_2
X_3
X_4
X_5
X

Fig. 5-2
Page 5-4

LINEAR FUNCTIONAL RELATIONSHIP OF TYPE FII

Fig 5-3
Page 5-5

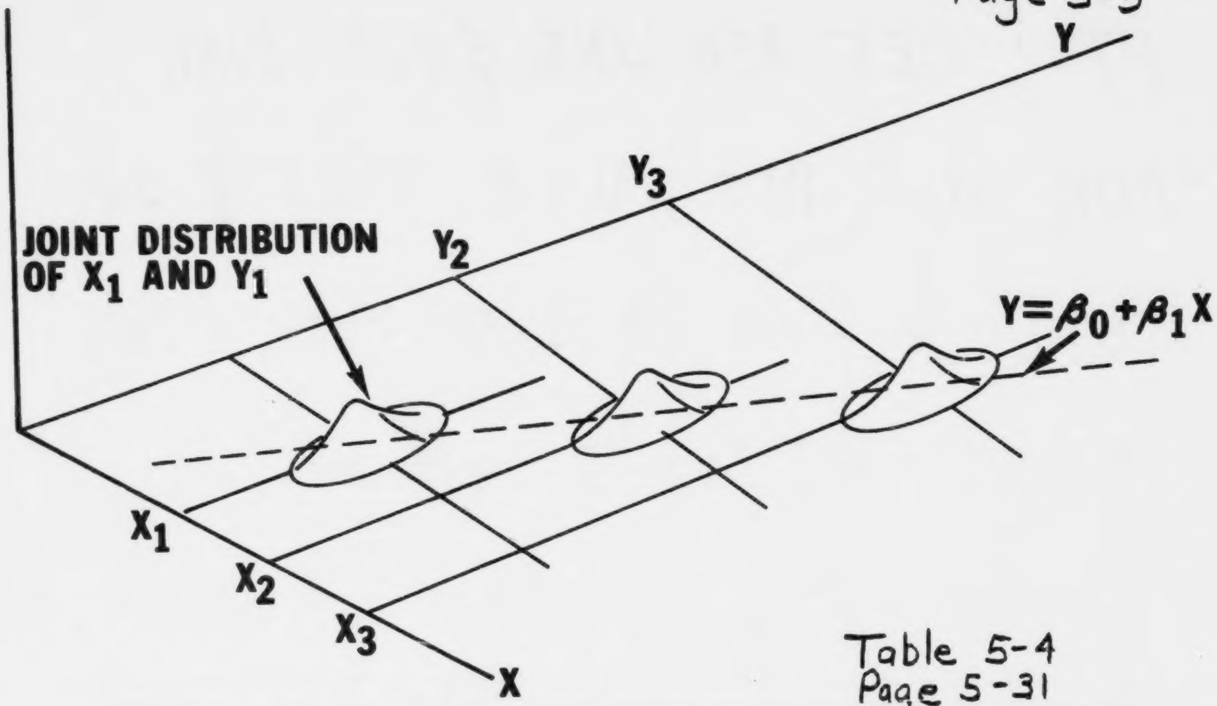

JOINT DISTRIBUTION OF X_1 AND Y_1
Y
Y_3
Y_2
$Y=\beta_0+\beta_1X$
X_1
X_2
X_3
X

Table 5-4
Page 5-31

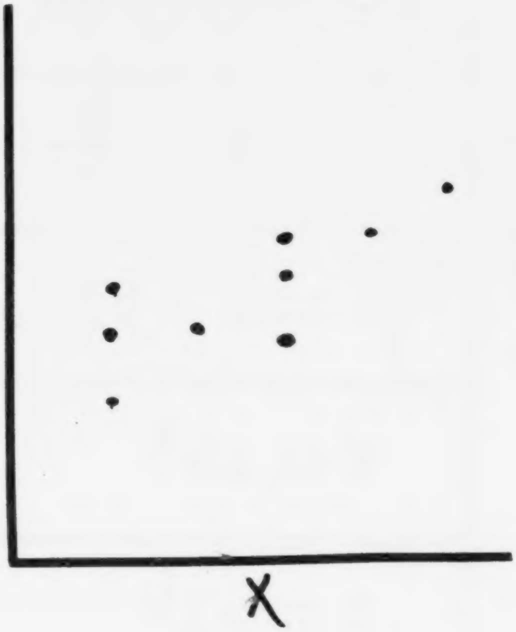
RELATIONSHIP	PLOT	
$Y = a + \frac{b}{X}$	Y	$\frac{1}{X}$
$\frac{1}{Y} = a + bX$	$\frac{1}{Y}$	X
$Y = ab^X$	$\log Y$	X
$Y = aX^b$	$\log Y$	$\log X$

OUTLIER PRESENT DELETE PT. DET. REG. LINE & TOL. LIM. FOR (n-1) PTS. USE T-TEST ON OUTLIER FOR PROB. OF SUCH A NO.

STAT. METHODS (6TH ED.)
SNEDECOR & COCHRAN P. 157

H_0: LIN. RELATIONSHIP (5-4.1.6)

MORE THAN ONE Y FOR ONE OR MORE VALUES OF X



H_0: LINEAR RELATIONSHIP
ART. 5–4.1.6

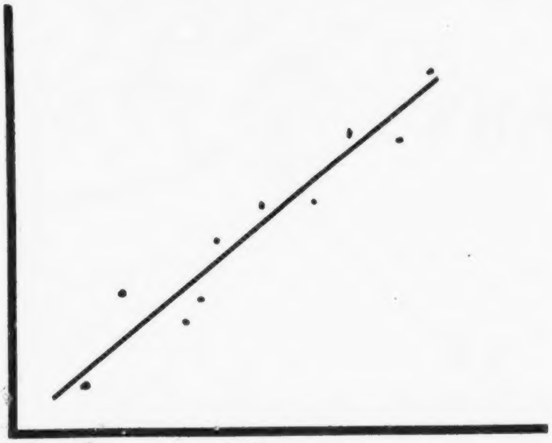
		REJECT H_0	DO NOT REJECT H_0
$H_0: \beta_1 = 0$ ART. 5–4.1.3	REJECT H_0	TRY $Y = \beta_0 + \beta_1 X + \beta_2 X^2$	USE $Y = b_0 + b_1 X$
	DO NOT REJECT H_0	TRY $Y = \beta_0 + \beta_1^* X + \beta_2 X^2$	USE $Y = \bar{Y}$

$^*\beta_1$ MAY NOT EQUAL ZERO WHEN SECOND DEGREE TERM IS INTRODUCED.

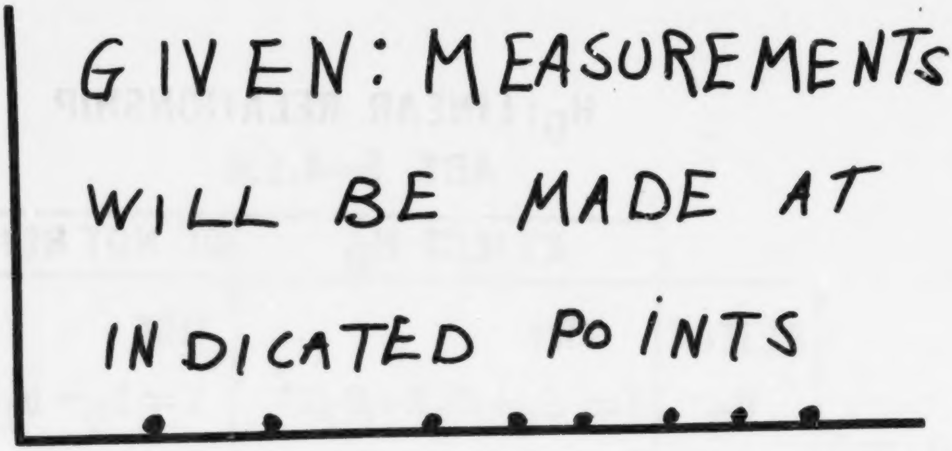
H_0: LIN. RELATIONSHIP

NON-PARAMETRIC - RUN TH.

CAN HAVE ONE Y FOR EVERY X



STAT. MANUAL, CROW, ET AL.
DOVER 1960 P. 165



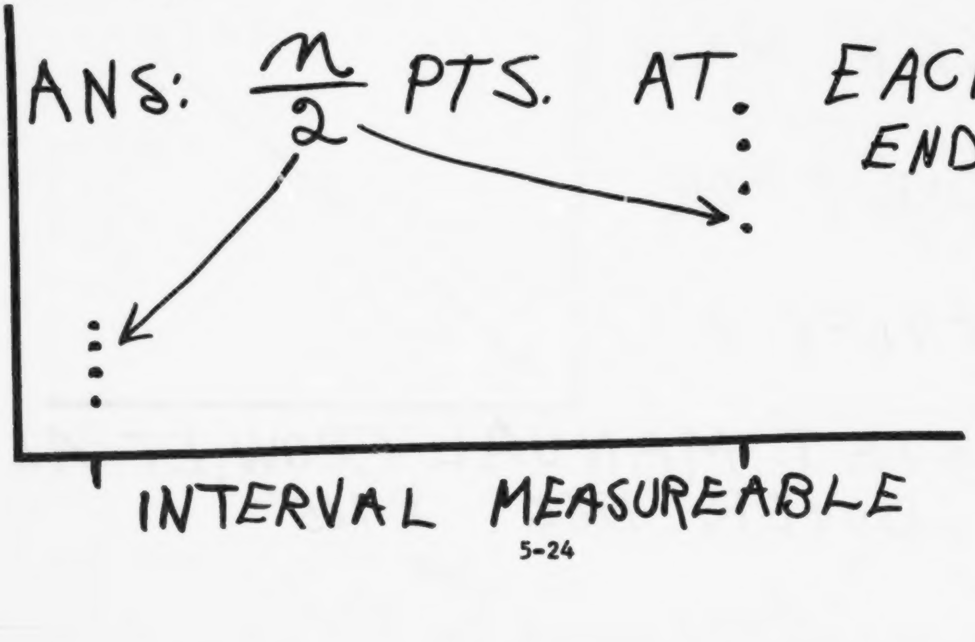
GIVEN: MEASUREMENTS WILL BE MADE AT INDICATED POINTS

RANDOMLY SELECT X'S

DON'T START AT ONE END AND TAKE IN ORDER.

GIVEN: KWN. LIN. RELTSHP.

DESIRED: BEST EST. FOR β_1

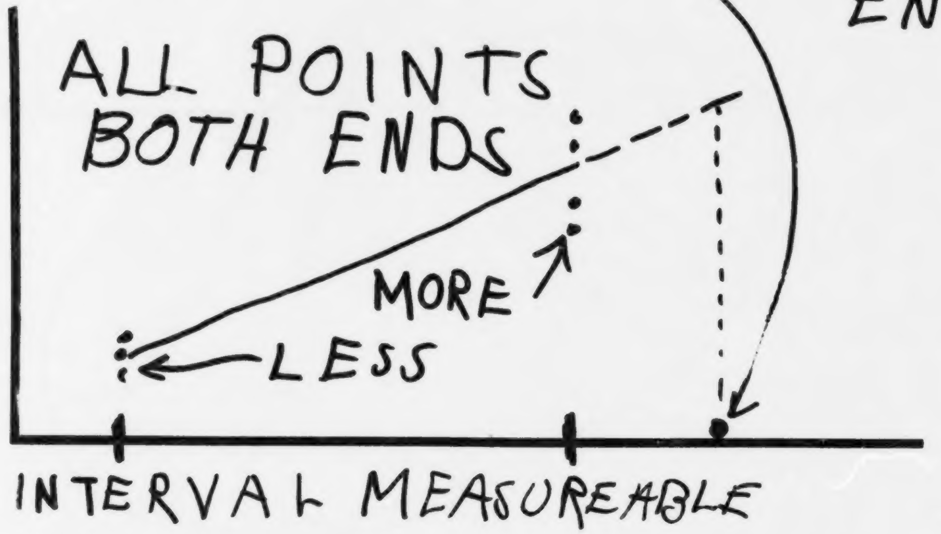


ANS: $\frac{n}{2}$ PTS. AT EACH END

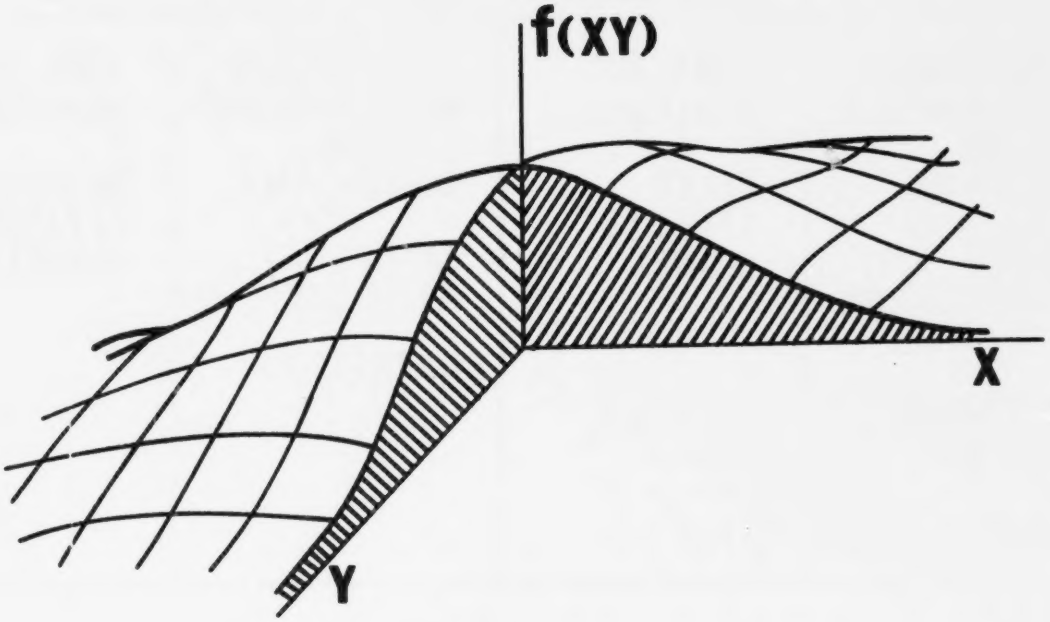
INTERVAL MEASUREABLE

KNOWN LIN. RELATIONSHIP

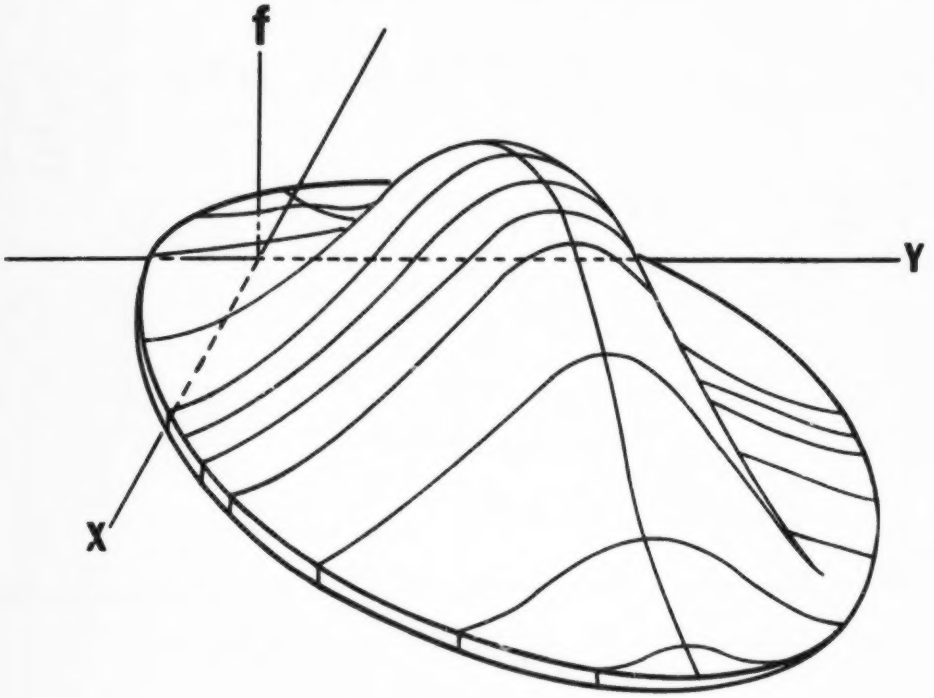
DESIRED: EXTRAPOLATE AT ONE END



BIVARIATE NORMAL DISTRIBUTION OR STATISTICAL RELATIONSHIP

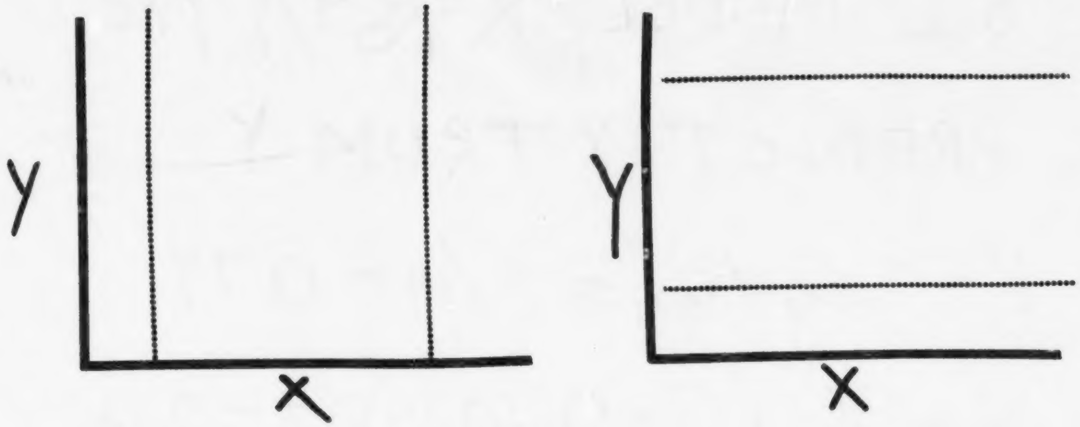


TWO-DIMENSIONAL NORMAL DISTRIBUTION OR STATISTICAL RELATIONSHIP



X FACTOR IS

CONTROLLED		UNCONTROLLED	
QUANTITATIVE	QUALITATIVE	QUANTITATIVE	QUALITATIVE
FUNCTIONAL RELATIONSHIP OR REG. ANA. MODEL FI NO ERROR IN X ONLY FII ERROR IN X & Y (NATRELLA)	ANA. OF VARIANCE FIXED EFFECTS MODEL	STATISTICAL RELATIONSHIP OR CORR. ANA. MODEL SI & SII (NATRELLA)	ANA. OF VARIANCE RANDOM EFFECTS MODEL



DATA RESTRICTED ON ONE AXIS ONLY → S II RELATIONSHIP - ART. 5-5.2

SI - MODEL SAME AS FI

Y = 3 + 1.1X ALSO SAME

CONF. INT. FOR $\bar{Y}$

TOL. LIMITS FOR IND. Y's

ETC.

SI- MODEL | $X = \beta_0' + \beta_1' Y + \varepsilon'$

PREDICT X FROM Y

$b_1' = S_{xy} / S_{yy} = 11/14 = 0.79$

$b_0' = \overline{X} - b_1' \overline{Y} = 0 - (0.79)3 = -2.4$

$X = -2.4 + 0.79Y$

$Y = 0 \rightarrow X = -2.4$ | $Y = 6 \rightarrow X = 2.34$

DATA BANK

$\Sigma X = 0$ $\quad \bar{X} = 0$ $\quad n = 4$

$\Sigma Y = 12$ $\quad \bar{Y} = 3$

$\Sigma XY = 11$ $\quad \Sigma X^2 = 10$ $\quad \Sigma Y^2 = 50$

$S_{xx} = 10$ $\quad S_{yy} = 14$

$S_{xy} = 11$ $\quad b_1 = 1.1$ $\quad b_0 = 3$

$Y = b_0 + b_1 X = 3 + 1.1X$

$s^2_Y = 0.95$ $\quad s_Y = 0.9747$ $\quad s_{b_1} = 0.3082$

$s_{b_0} = 0.4873$ $\quad t_{.975}[df = 2] = 4.303$

PAGE T-5

SI- MODEL ¦ $X=\beta'_0+\beta'_1 Y+\varepsilon'$

PREDICT X FROM Y

$$b'_1 = S_{xy}/S_{yy} = 11/14 = 0.79$$

$$b'_0 = \overline{X} - b'_1\overline{Y} = 0-(0.79)3 = -2.4$$

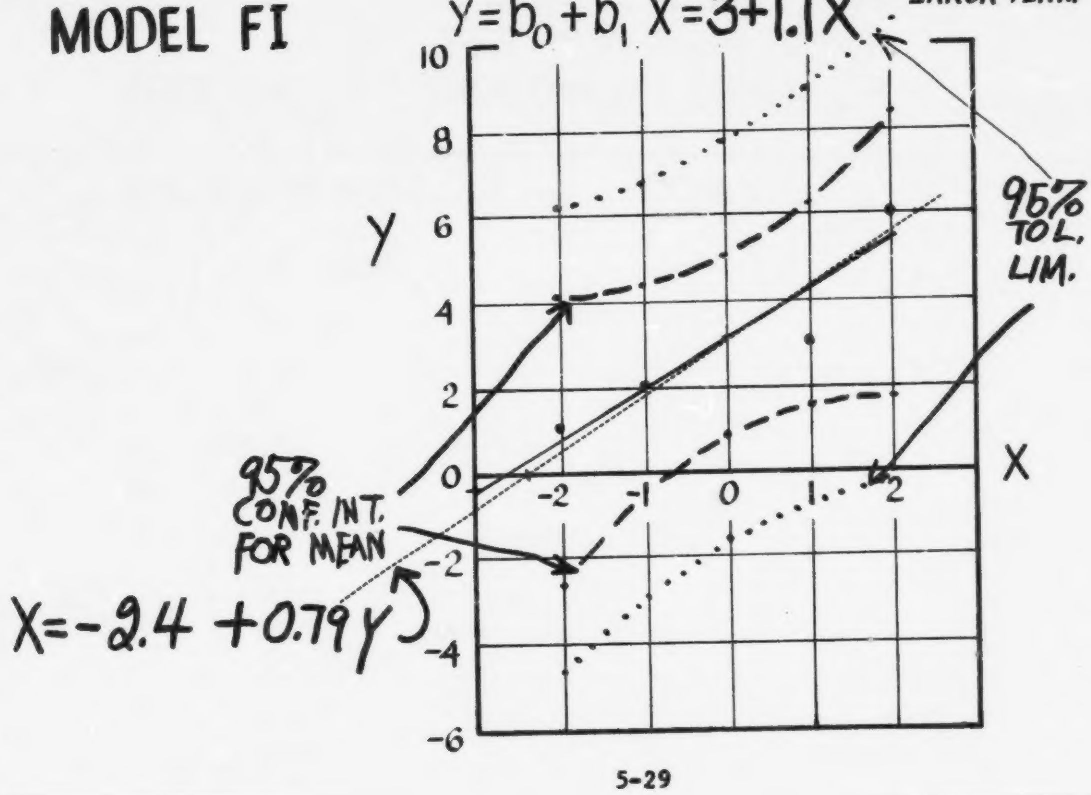
$$X = -2.4 + 0.79Y$$

$Y=0 \rightarrow X=-2.4$ ¦ $Y=6 \rightarrow X=2.34$

$Y=\beta_0+\beta_1 X+\varepsilon$ (MODEL)

$Y=b_0+b_1 X=3+1.1X$

MODEL FI



CORRELATION COEFF. $\equiv \rho$

LIMITS $-1 \leq \rho \leq 1$ ¦ SAMP. EST. $= r$

$$r = S_{xy} / \sqrt{S_{xx} S_{yy}} = 11 / \sqrt{10(14)} = 0.9297$$

DATA BANK

$\Sigma X = 0$ $\bar{X} = 0$ $n = 4$

$\Sigma Y = 12$ $\bar{Y} = 3$

$\Sigma XY = 11$ $\Sigma X^2 = 10$ $\Sigma Y^2 = 50$

$S_{xx} = 10$ $S_{yy} = 14$

$S_{xy} = 11$ $b_1 = 1.1$ $b_0 = 3$

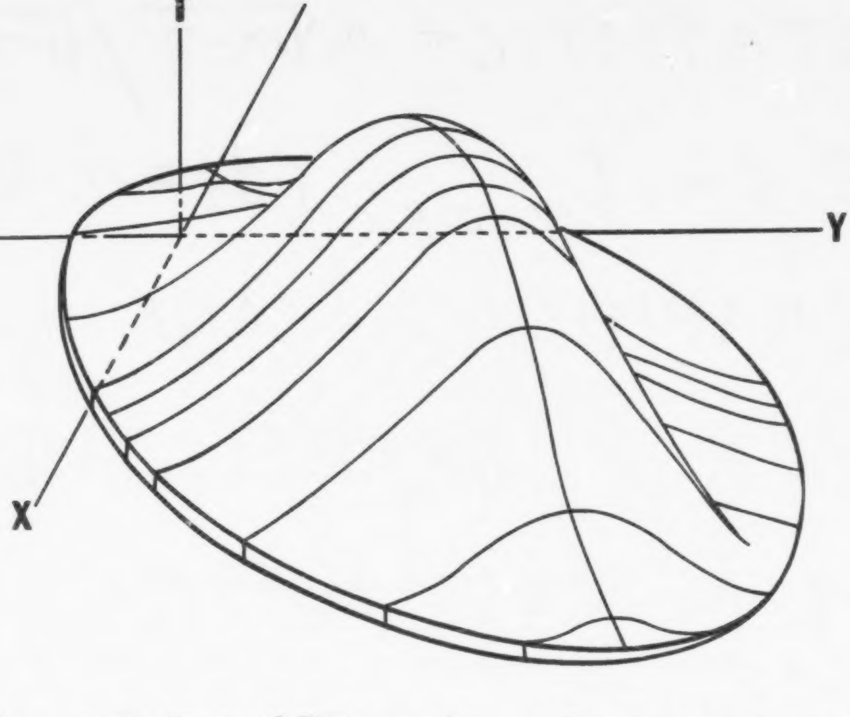
$Y = b_0 + b_1 X = 3 + 1.1X$

$s^2_Y = 0.95$ $s_Y = 0.9747$ $s_{b_1} = 0.3082$

$s_{b_0} = 0.4873$ $t_{.975}[df = 2] = 4.303$

PAGE T-5

TWO-DIMENSIONAL NORMAL DISTRIBUTION OR STATISTICAL RELATIONSHIP



CORRELATION COEFF. $\equiv \rho$

LIMITS $-1 \le \rho \le 1$ | SAMP. EST. $= r$

$$r = S_{xy} / \sqrt{S_{xx} S_{yy}} = 11 / \sqrt{10(14)} = 0.9297$$

$r^2 \equiv$ COEF. OF DETERMINATION

PROP. OF VAR. OF Y GIVEN BY X

$$r^2 = 0.9297^2 \cong 0.864 \text{ OR } 86.4\%$$

$H_0: \rho = 0$ ¦ $H_A: \rho \neq 0$

STATISTIC $= r\sqrt{n-2} / \sqrt{1-r^2}$

C.V. $= t_{1-0.5\alpha}$ $(df = n-2)$

GRAPHICAL CONF. INT.

FOR ρ ¦ NATRELLA T-31

$H_0: \rho = \text{CONST.}$ ¦ $H_0: \rho_1 = \rho_2 = \cdots = \rho_k$

GIVEN: CHOICE OF TWO

VAR. FOR PRED. THE THIRD

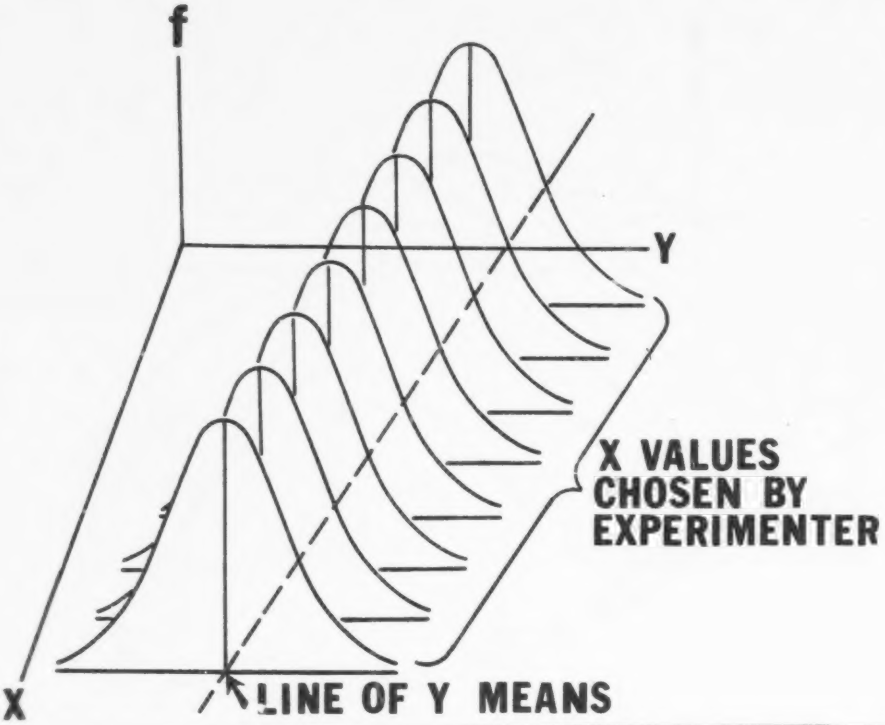
$Y = b_0 + b_1 X$ OR $Y = b_0^* + b_1^* Z$

STAT. INFERENCE, 1953,

WALKER, LEV ¦ HOLT

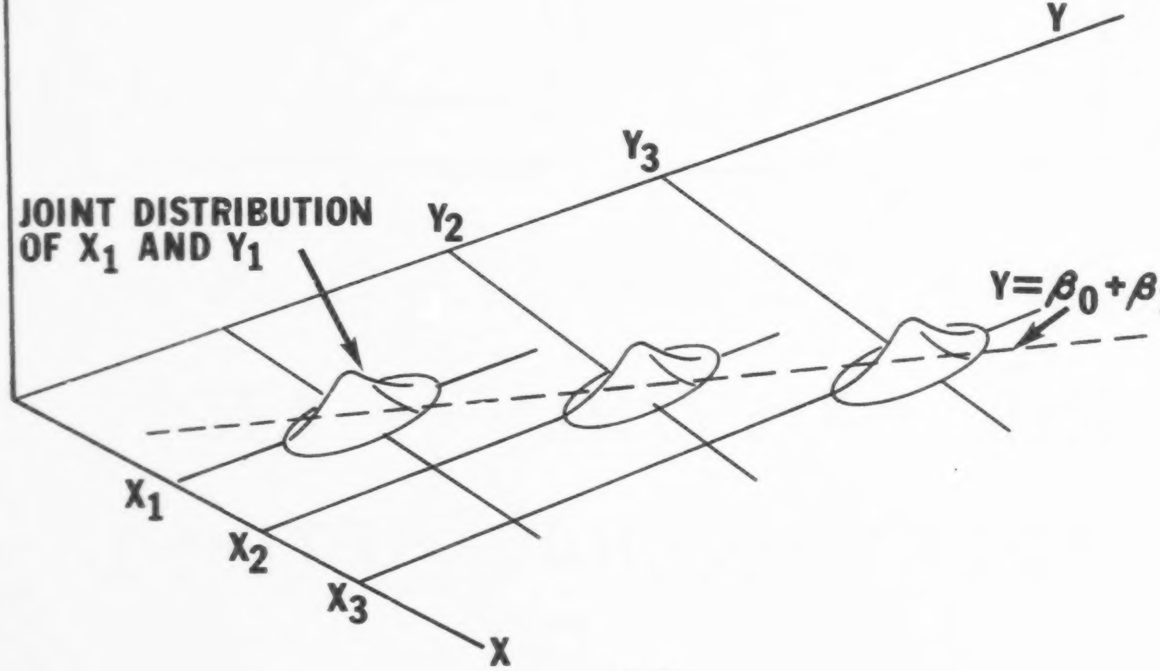
P. 256

FUNCTIONAL RELATIONSHIP FI

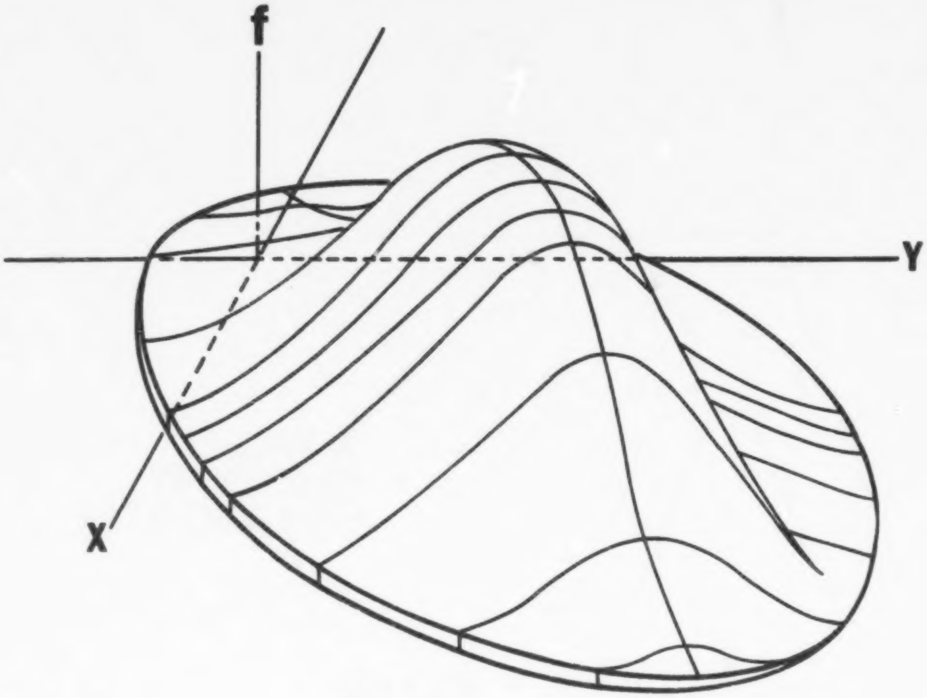

f
Y
X VALUES
CHOSEN BY
EXPERIMENTER
X
LINE OF Y MEANS

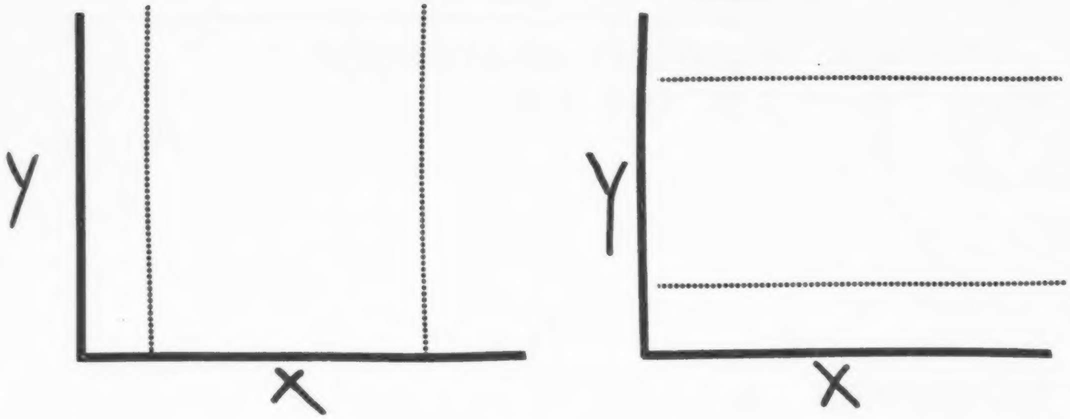
LINEAR FUNCTIONAL RELATIONSHIP OF TYPE FII

Fig 5-3
Page 5-5


Y
Y_3
Y_2
JOINT DISTRIBUTION
OF X_1 AND Y_1
$Y=\beta_0+\beta_1 X$
X_1
X_2
X_3
X

TWO-DIMENSIONAL NORMAL DISTRIBUTION OR STATISTICAL RELATIONSHIP



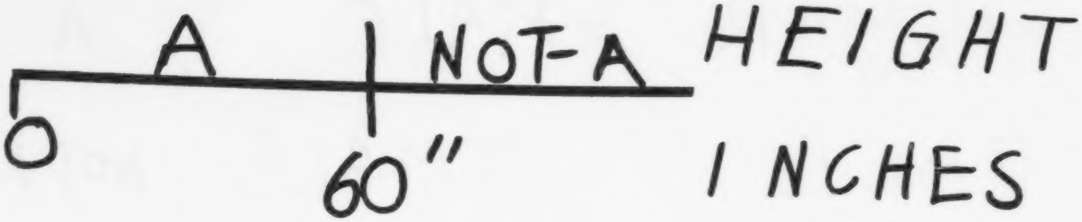


DATA RESTRICTED ON ONE AXIS ONLY → S II RELATIONSHIP - ART. 5-5.2

	Bi Pop.	DES.
HUMAN RACE	MALE	A
	FEMALE	NOT-A

COIN FACE UP	HEAD	A
	TAIL	NOT-A

FACES UP ON A DIE				
1	A		A	
2		NOT-A		NOT-A
3		"	A	
4		"		"
5		"	A	
6		"		"
	Bi.	Pop.	Bi.	Pop.


A
NOT-A
HEIGHT
0
60"
INCHES

A ≡ PEOPLE 60" OR LESS

NOT-A = " OVER 60"

TAKE n IND. TRIALS

FROM A BIN. POP. p = CONST.

$P(\text{A-SIN. TR.}) = p$; $P(\text{NOT-A, SIN. TR.}) = q$

$$P\begin{pmatrix}\text{A OCCURS}\\ \text{X TIMES}\end{pmatrix} = \frac{n!}{(n-x)!\,x!}\,p^{x}q^{n-x}$$

WHERE $n! = 1 \cdot 2 \cdots n$

DEFINITION $0! = 1$

$$P\begin{pmatrix} X;\ A\text{'}A \\ \text{OUT OF } n \end{pmatrix} = \frac{n!}{(n-X)!\,X!}\,p^{X}\,q^{n-X}$$

TOSS 10 DIE $\;|\;$ $n=10$ $\;|\;$ PROB. OF 6 ACES $\;|\;$ $X=6$

$$P(\text{ACE, SIN. TR.}) = \frac{1}{6} = p \;|\; q = 1-p = \frac{5}{6}$$

$$P\begin{pmatrix} 6,\ \text{ACE'S} \\ \text{OUT OF } 10 \end{pmatrix} = \frac{10!}{4!\;6!}\left(\frac{1}{6}\right)^{6}\left(\frac{5}{6}\right)^{4}$$

PARAMETERS BIN. DISTB.

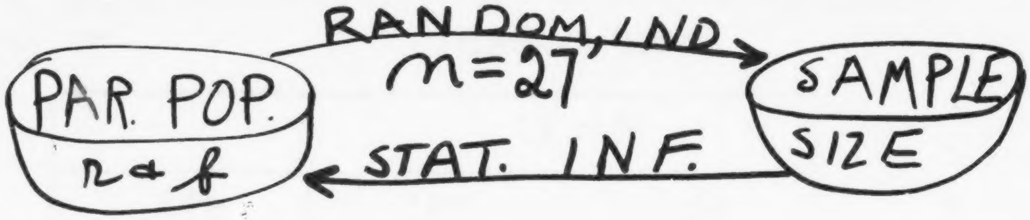
$$n=100 \;|\; p=\frac{1}{2}=q \;|\; \mu_{B,C} = np = \frac{100}{2} = 50$$

$$\sigma_{B,C} = \sqrt{npq} = \sqrt{100\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 5$$

COUNTS

$$\mu_{B,R} = p \;|\; \sigma_{B,R} = \sqrt{\frac{pq}{n}}$$

RATIOS


RANDOM, IND
$n=27$
PAR. POP.
r & p
SAMPLE
SIZE
STAT. INF.

$r =$ NO. OF RED BALLS $= 11$

p (EST. P; PAR. POP. RATIO) $= r/n = 11/27 = 0.41$

SINCE $10/27 = .37$; $.37 < P < 0.41$ IMP.

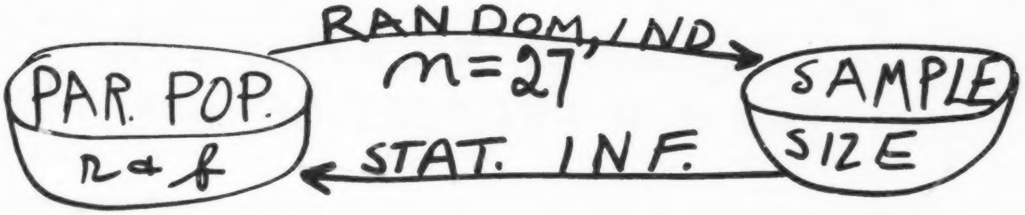
TWO-SIDED CONF. INTERVALS FOR PARENT POP. PROPORTION P

NATURE	SAMPLE	ART.	METHOD
→ EXACT	$n \leq 30$	7-3.1.1	TABLE A-22
EXACT	$n > 50$	7-3.1.2	TABLE A-24
APPROX.	$n > 30$	7-3.1.3	FORMULA

CONFIDENCE LIMITS FOR PROPORTION (TWO-SIDED)

Table A-22
T-40

	n = 27		
r	**90%**	**95%**	**99%**
0			
1			
⋮			
11	.239 .593	.223 .598	
⋮			
27			

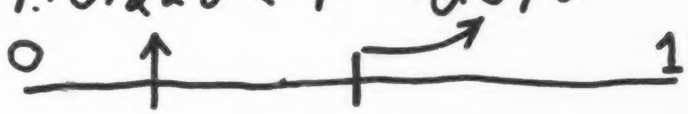

RANDOM, IND
$n = 27$
PAR. POP.
r & f
SAMPLE
SIZE
STAT. INF.

r = NO. OF RED BALLS = 11

$\hat{p}$ (EST. P; PAR. POP. RATIO) $= r/n = 11/27 = 0.41$

SINCE $10/27 = .37$; $.37 < P < 0.41$ IMP.

95% CONF. INT: $0.223 < P < 0.598$



TWO-SIDED CONF. INTERVALS FOR PARENT POP. PROPORTION P

NATURE	SAMPLE	ART.	METHOD
EXACT	$n \leq 30$	7-3.1.1	TABLE A-22
EXACT	$n > 50$	7-3.1.2	TABLE A-24
→ APPROX.	$n > 30$	7-3.1.3	FORMULA

BIN $\rightarrow$ NOR: $n \geq 30$; $np \geq 5$ | $nq \geq 5$

$(p - \mu_{B,R}) / \sigma_{B,R} = Z$ | $100(1-\alpha)\%$ TWO-SIDED CONF. INT. : $n = 100$; $p = .4$ | $\alpha = 0.05$

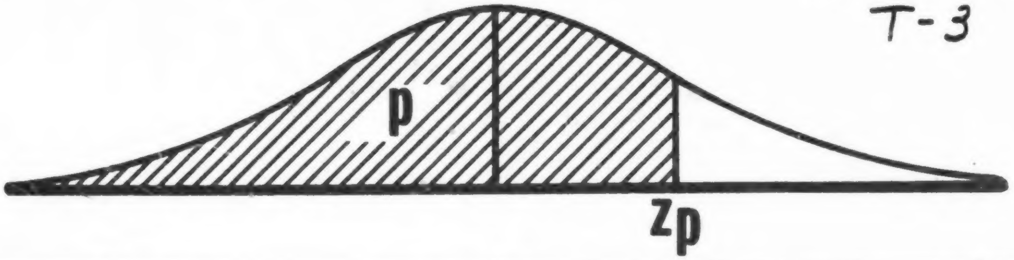
$$p - Z_{1-0.5\alpha}\sqrt{pq/n} < P < p + Z_{1-0.5\alpha}\sqrt{pq/n}$$

$$0.4 - 1.96\sqrt{.4(.6)/100} < P < .4 + 1.96\sqrt{.4(.6)/100}$$

$.4 - 0.096 < P < .4 + 0.096$ | $0.304 < P < .496$

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090

ONE-SIDED CONF. INTERVAL FOR P

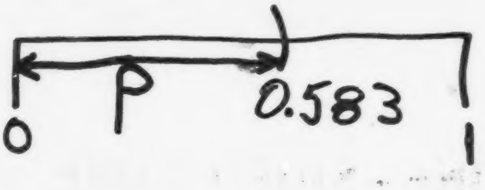
NATURE	SAMPLE	ART.	METHOD
→ EXACT	$n \leq 30$	7–3.2.1	TABLE A–23
EXACT	$n > 50$	7–3.2.2	TABLE A–24
APPROX.	$n > 30$	7–3.2.3	FORMULA

CONFIDENCE LIMITS FOR PROPORTION (ONE-SIDED)

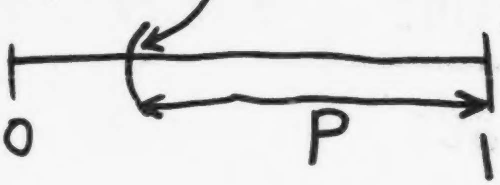
Table A-23
T-43

	$n = 27$		
r	90%	95%	99%
0			
⋮			
11	.549	.583	$.645^+$
⋮			
16		.752	
⋮			
26			

$n=27$ ⁞ $r=11$ ⁞



$n-r=16$ ⁞ $1-0.752=0.248$

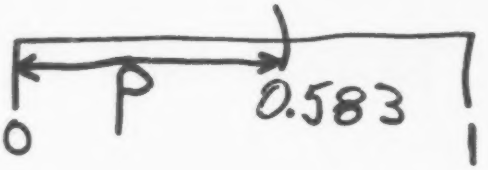


CONFIDENCE LIMITS FOR PROPORTION
(ONE-SIDED)

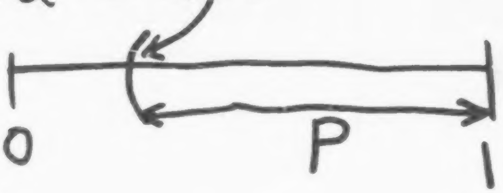
Table A-23
T-43

r	n = 27		
	90%	95%	99%
0			
⋮			
11	.549	.583	.645^{+}
⋮			
16		.752	
⋮			
26			

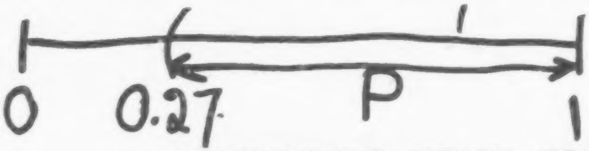
$n=27$ ¦ $r=11$ ¦



$n-r=16$ ¦ $1-0.752=0.248$

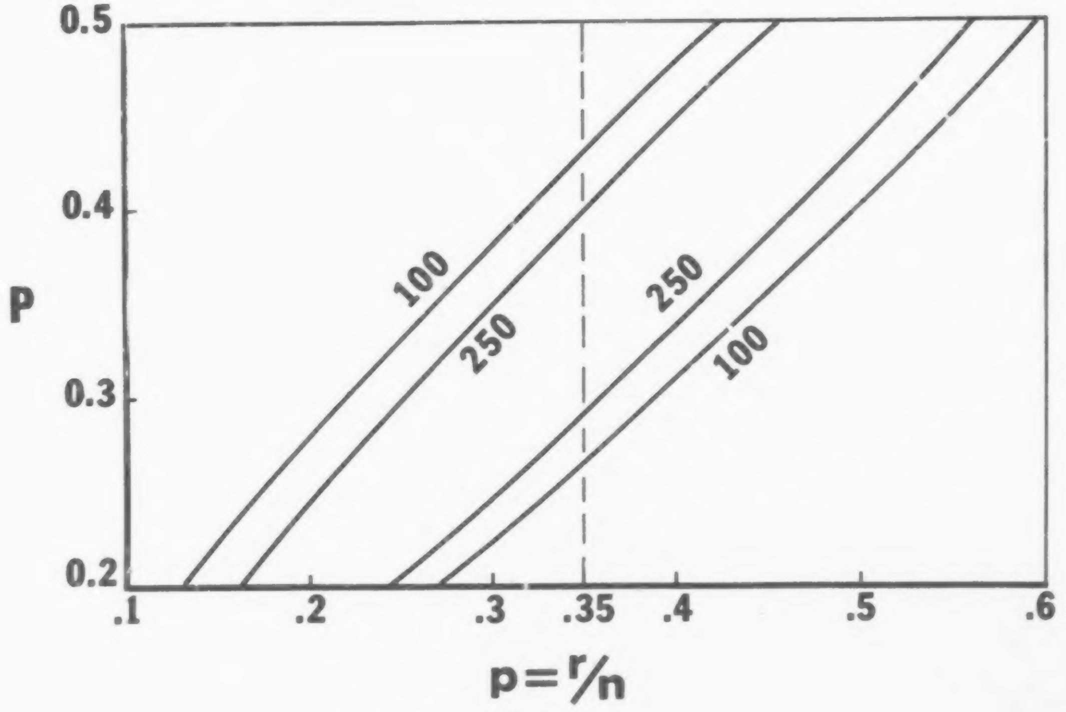


$r/n=0.35$ ¦ $n=100$ ¦ 95% CONF. INT



CONFIDENCE BELTS FOR PROPORTIONS
(CONFIDENCE COEFFICIENT 0.90)

Table A-24
T-45



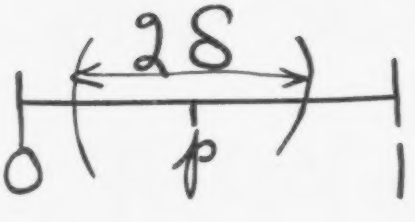
SAMPLE SIZE DETERMINATION

TWO-SIDED CONFIDENCE INTERVAL

NATURE	SAMPLE	ART.	METHOD
EXACT	$n > 50$	7–4.1.1	TABLE A–24
APPROX.	$n > 30$	7–4.1.2	FORMULA

ONE-SIDED CONFIDENCE INTERVAL

EXACT	$n > 30$	7–4.2	FORMULA

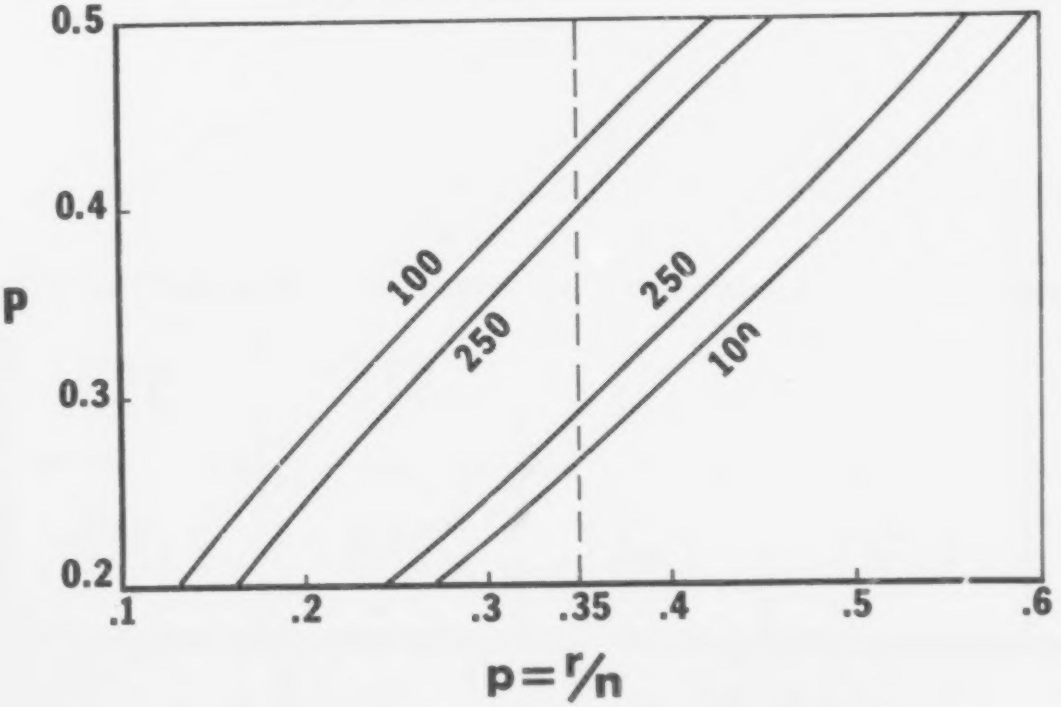


$$n = \frac{Z^2_{1-0.5\alpha} P(1-P)}{\delta^2}$$

CASE	P	CHOICE OF P
I	UNK.	0.5
II	0 TO 0.5	LG. EST. OF P
III	0.5 TO 1	SM. EST. OF P

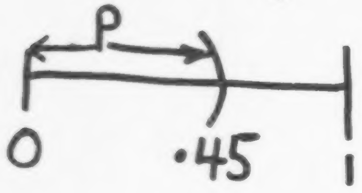
CONFIDENCE BELTS FOR PROPORTIONS
(CONFIDENCE COEFFICIENT 0.90)

Table A-24
T-45



$\delta = 0.05$ ¦ $\alpha = 0.05$ ¦ P(UNK) → $P = 0.5$

$$n = \frac{Z_{0.975}^2 P(1-P)}{\delta^2} = \frac{1.96^2}{4(.05)^2} = 384.2 = 385$$



$$n = \frac{1.96^2 (.45)\, 0.55}{(0.05)^2}$$

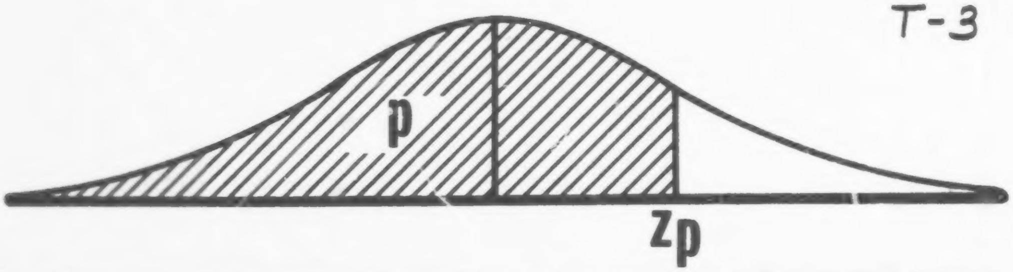
1-SIDED

LIMITS

$$n = \frac{Z_{1-\alpha}^2 P(1-P)}{\delta^2}$$

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
Z_p	1.282	1.645	1.960	3.090

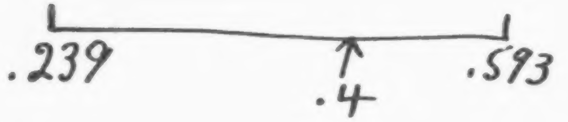
ONE SAMPLE PROBLEM

TEST	CONDITIONS
H_0 : $P =$ CONSTANT H_A : $P \neq$ CONSTANT	1. $N \leq 30$ 2. $N > 30$
H_0 : $P \leq$ CONSTANT H_A : $P >$ CONSTANT	1. $N \leq 30$ 2. $N > 30$

$P \equiv$ PARENT POPULATION PROPORTION

$H_0: P = 0.4$, $H_A: P \neq 0.4$, $\alpha = .1$

$n=27$, $r=11$, $1-\alpha=0.9$, A-22, PT-40



90% 2-S
CONF. INT.

IF CONST. IS WITHIN LIMITS DO NOT REJ. H_0, OTHERWISE REJ. H_0

DEC. DO NOT REJ. H_0

CONFIDENCE LIMITS FOR PROPORTION
(TWO-SIDED)

Table A-22
T-40

r	n = 27		
	90%	95%	99%
0			
1			
⋮			
11	.239 .593		
⋮			
27			

n = 1(1)30

$H_0: P=0.4,\ H_A: P\neq 0.4,\ \alpha=0.1=\beta$

$\delta \equiv$ ABS. MAG. OF DIFF. TO BE DET.

$\delta=0.1,\ P=0.4\pm\delta$ (VALUE CLO.
TO 0.5), $P=0.5$ OR 0.3 $\therefore P=0.5$

$$d=\left|2\sin^{-1}\sqrt{P}-2\sin^{-1}\sqrt{.4}\right|=\left|1.57-1.37\right|=0.2$$

$$m=\left(\frac{Z_{1-0.5\alpha}+Z_{1-\beta}}{d}\right)^2=\left(\frac{1.645+1.282}{0.2}\right)^2=214$$

TABLE OF ARC SINE ... PROPORTIONS

$\theta = 2\ \text{ARC SIN}\sqrt{P}$

Table A-27
T-54

P	θ	P	θ	P	θ
		0.4	1.37		
		0.5	1.57		

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

P	.9	.95	.975	.999
z_p	1.282	1.645	1.960	3.090

Table A-2
T-3

$H_0: P = 0.4$, $H_A: P \neq 0.4$, $\alpha = 0.1 = \beta$

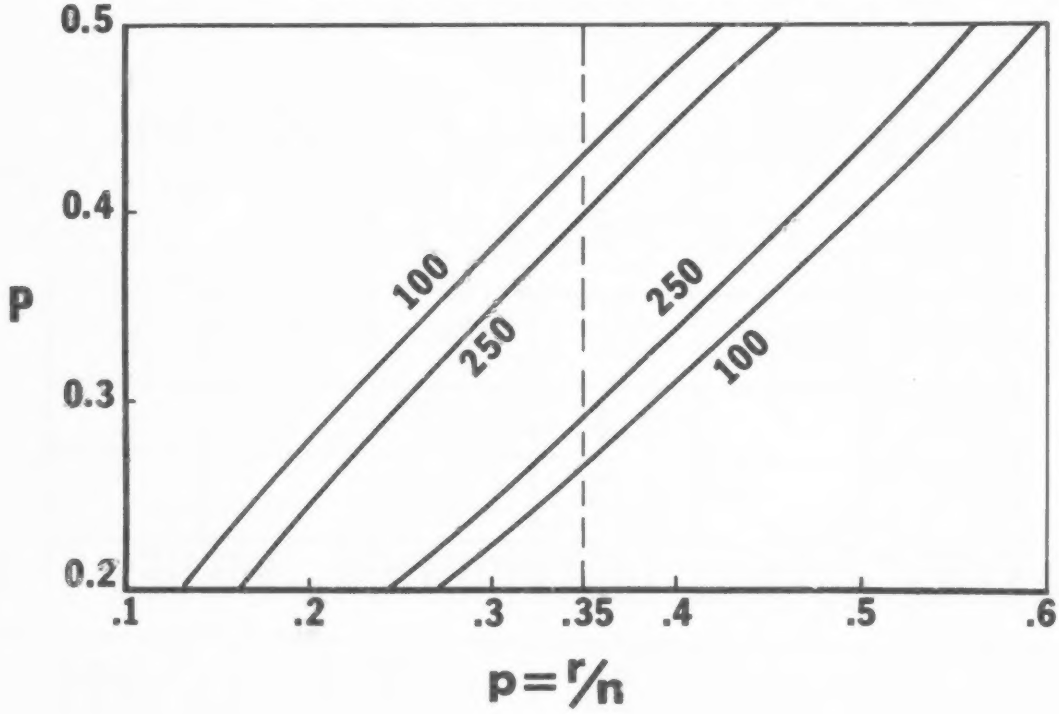
$n = 214$, $r = 75$, $p = r/n = 75/214 = 0.35$

90% CONF. INT., .29 $\overset{\uparrow}{.4}$ .41

DEC. DO NOT REJ. H_0

CONFIDENCE BELTS FOR PROPORTIONS
(CONFIDENCE COEFFICIENT 0.90)

Table A-24
T-45



ONE SAMPLE PROBLEM

TEST	CONDITIONS
$H_0 : P = \text{CONSTANT}$ $H_A : P \neq \text{CONSTANT}$	1. $N \leq 30$ 2. $N > 30$
$H_0 : P \leq \text{CONSTANT}$ $H_A : P > \text{CONSTANT}$	1. $N \leq 30$ 2. $N > 30$

$P \equiv$ PARENT POPULATION PROPORTION

ONE SAMPLE PROBLEM

TEST	CONDITIONS
$H_0 : P \geq \text{CONSTANT}$ $H_A : P < \text{CONSTANT}$	1. $N \leq 30$ 2. $N > 30$

$P \equiv$ PARENT POPULATION PROPORTION

TWO SAMPLE PROBLEM

TEST	CONDITIONS
$H_0 : P_A = P_B$ $H_A : P_A \neq P_B$	1. $N_A = N_B$ 2. $N_A \neq N_B$ (BOTH LESS THAN 20) 3. $N_A \neq N_B$ (LARGE SAMPLES)

$P_A \equiv$ PROPORTION FOR A POPULATION

$N_A \equiv$ SAMPLE SIZE FROM A POPULATION

$H_0: P_A = P_B$, $H_A: P_A \neq P_B$, $\alpha = 0.05$

	SUB.	NOT
A	2	18
B	8	12

$a_1 \equiv$ SM. TABLE ENT. $= 2$

$a_2 \equiv$ OTHER COL. VALUE $= 8$

FIND A_1 & A_2 TABLE A-28, WHERE $a_1 = A_1$

$A_2 = 9$ IF $a_2 \geq A_2$ REJ. H_0 AR RR SA.

8 9

DEC. DO NOT REJ. H_0 SINCE $8 < 9$

MINIMUM CONTRAST...

Table A-28
T-56

5% LEVEL, TWO-SIDED (ALSO 1%) 2.5% LEVEL, ONE-SIDED (ALSO 0.5%)	
SAMPLE SIZE $n_A = n_B$	**A_1, A_2**
⋮	
20	0,5 1,7 2,9 3,10 4,11 5,13 6,14
⋮	

$n_A = n_B = 1\ (1)\ 20\ (10)\ 100\ (50)\ 200\ (100)\ 500$

$H_0: P_A = P_B,\ H_A: P_A \neq P_B,\ \alpha = 0.05,\ \beta = 0.1$

$P' \equiv$ EST OF P_A OR P_B (CONS. NEAR 0.5) = 0.4

$\delta \equiv$ ABS. MAG. OF DIFF. TO BE DET. = 0.08

$P'' = P' \pm \delta$ (CLO. TO 0.5) = 0.4 ± .08 = 0.48

$d = |2\sin^{-1}\sqrt{P'} - 2\sin^{-1}\sqrt{P''}| = |1.37 - 1.53| = 0.16$

$$n = 2\left(\frac{Z_{1-0.5\alpha} + Z_{1-\beta}}{d}\right)^2 = 2\left(\frac{1.96 + 1.282}{0.16}\right)^2 = 822$$

TABLE OF ARC SINE ... PROPORTIONS

$\theta = 2\ \text{ARC SIN}\sqrt{P}$

Table A-27
T-54

P	θ	P	θ	P	θ
		0.4	1.37		
		0.48	1.53		

A-2
T-3

CUMULATIVE NORMAL DISTRIBUTION—VALUES OF z_p

P	.9	.95	.975	.999
z_p	1.282	1.645	1.960	3.090

	SUB. STAND.	ACCEPT.	
A	$r_A = 181$	$s_A = 641$	$n_A = 822$
B	$r_B = 160$	$s_B = 662$	$n_B = 822$
	$r = 341$	$s = 1303$	$n = 1644$

$$ST_o = \frac{n\left(\left|r_A s_B - r_B s_A\right| - 0.5n\right)^2}{n_A n_B \, r \, s}$$

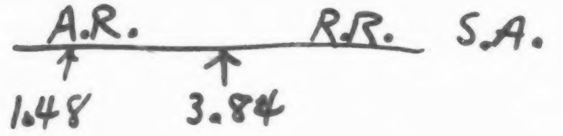
$$= \frac{1644\left(\left|181(662) - 160(641)\right| - 0.5(1644)\right)^2}{822(822) \; 341 \; (1303)}$$

$$= 1.48$$

$H_0: P_A = P_B$, $H_A: P_A \neq P_B$, $\alpha = .05$, $\beta = .1$

ST. = 1.48   DISTB. $\chi^2 (df=1)$

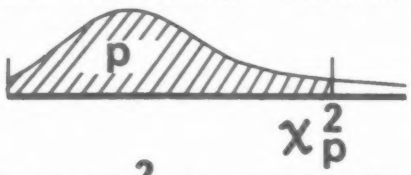
CRIT. VALUE $= \chi^2_{1-\alpha} (df=1) = \chi^2_{.95} = 3.84$



DEC. DO NOT REJ. H_o

PERCENTILES OF THE χ^2 DISTRIBUTION

Table A-3
T-4



THE VALUES OF χ^2_p CORRESPONDING TO p

df	...	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.99}$
1			3.84	6.63
4			9.49	
⋮				
10		15.99		

TWO SAMPLE PROBLEM

TEST	CONDITIONS
$H_0: P_A = P_B$ $H_A: P_A > P_B$	1. $N_A = N_B$ 2. $N_A \neq N_B$ (BOTH LESS THAN 20) 3. $N_A \neq N_B$ (LARGE SAMPLES)

$P_A \equiv$ PROPORTION FOR A POPULATION

$N_A \equiv$ SAMPLE SIZE FROM A POPULATION

MINIMUM CONTRAST...

Table A-28
T-56

5% LEVEL, TWO-SIDED (ALSO 1%) 2.5% LEVEL, ONE-SIDED (ALSO 0.5%)	
SAMPLE SIZE $n_A = n_B$	A_1, A_2
⋮	
20	0,5 1,7 2,9 3,10 4,11 5,13 6,14
⋮	

$n_A = n_B = 1\,(1)\,20\,(10)\,100\,(50)\,200\,(100)\,500$

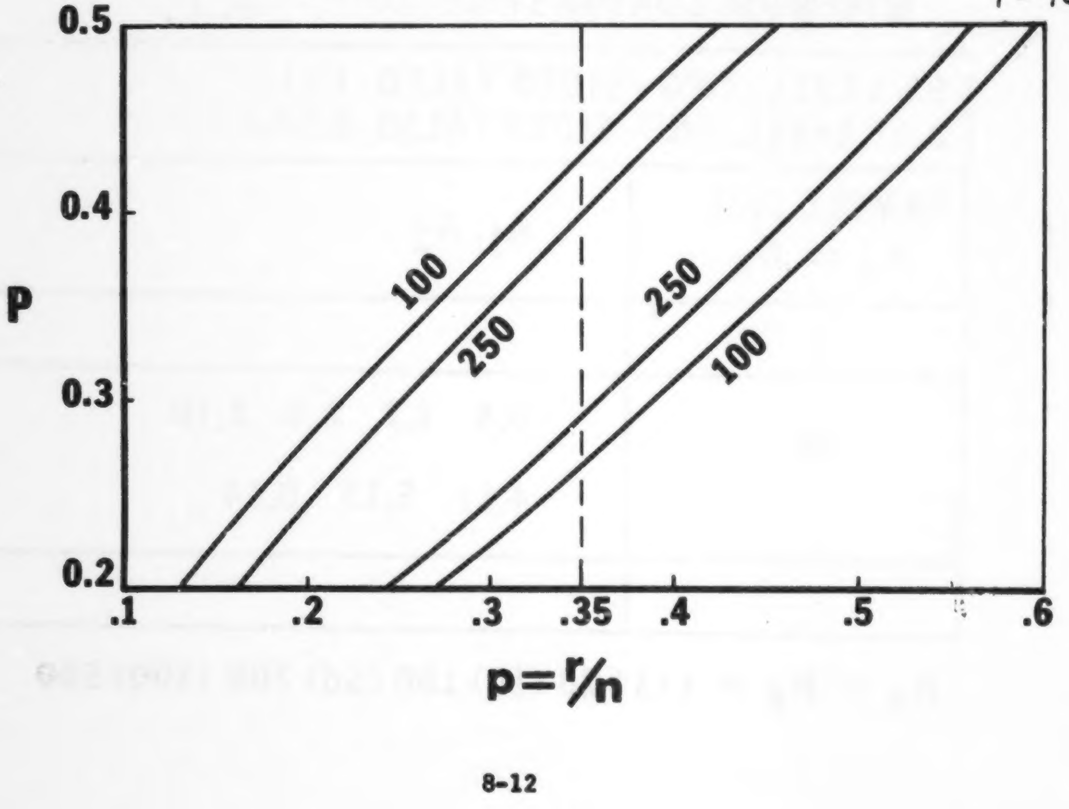
ONE SAMPLE PROBLEM

TEST	CONDITIONS
H_0 : $P \geq$ CONSTANT H_A : $P <$ CONSTANT	1. $N \leq 30$ 2. $N > 30$

$P \equiv$ PARENT POPULATION PROPORTION

CONFIDENCE BELTS FOR PROPORTIONS
(CONFIDENCE COEFFICIENT 0.90)

Table A-24
T-45



	CATEGORIES OR REJECTION (DISCRETE)			
	SAND	DROPPED	BROKEN	OTHER
NIGHT SHIFT				

	CATEGORIES OR MEASUREMENT (CONTINUOUS)			
	$-\infty$ TO -1	-1 TO 0	0 TO 1	1 TO ∞
ANNUAL PROD				

	CATEGORIES OR REJECTION (DISCRETE)			
	SAND	DROPPED	BROKEN	OTHER
WEEK 1				
WEEK 2				
WEEK 3				

		FIRST CATEGORY OR TYPE OF SUCCESS		
		A	B	C
SECOND CATEGORY OR EDUCATION LEVEL	GRADE SCHOOL			
	HIGH SCHOOL			
	COLLEGE			
	GRADUATE			

$X \equiv$ CELL COUNTS OR FREQ.

IF X IN CELLS ARE POISSON DISTB. $\mu = \sigma^2$

$$\frac{X-\mu}{\sigma} \xrightarrow{\text{TRANS}} \text{NOR., IF } \mu \geq 5$$

$$\left(\frac{X-\mu}{\sigma}\right)^2 = \frac{(X-\mu)^2}{\sigma^2} = \frac{(X-\mu)^2}{\mu} = \left\{\begin{matrix}\text{SQ.}\\ \text{NOR.}\end{matrix}\right.$$

$$ST = \sum_{\text{ALL}}^{\text{CELLS}} \frac{(X-\mu)^2}{\mu} \rightarrow \chi^2 \text{ DISTB}$$

df = # OF CELLS −

OF CONSTRAINTS

	RANDOM DIGITS					
	0, 1	2, 3	4, 5	6, 7	8, 9	TOTAL
O	30	40	33	47	50	200=n

H_0: $P_i = 1/5$ FOR EVERY i

H_A: NOT ALL $P_i = 1/5$

$\alpha = 0.05$

	RANDOM DIGITS					
	0, 1	2, 3	4, 5	6, 7	8, 9	TOTAL
O	30	40	33	47	50	200 = n
E	40	40	40	40	40	200

$\mu = n P_i = 40 = \text{AVE} = E \quad \text{MUST} \overset{\uparrow}{=} n$

$$ST = \sum \frac{(X-\mu)^2}{\mu} = \sum \frac{(O-E)^2}{E}$$

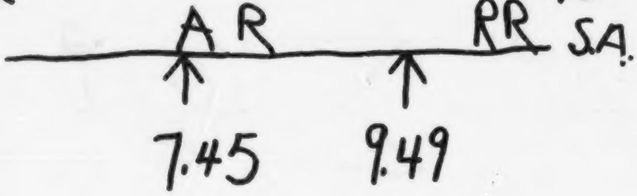
$$= (30-40)^2/40 + \cdots + (50-40)^2/40 = 7.45$$

H_0 $P_i = 1/5$ FOR EVERY i

H_A: NOT ALL $P_i = 1/5$

$\alpha = 0.05$ $df = 5 - 1 = 4$

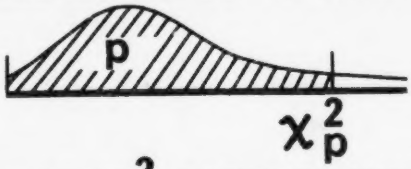
$\chi^2 = 7.45$ $CV = \chi^2_{.95}(df = 4) = 9.49$



DEC. DO NOT REJ H_0

PERCENTILES OF THE χ^2 DISTRIBUTION

Table A-3
T-4



THE VALUES OF χ^2_p CORRESPONDING TO p

df	...	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.99}$
1			3.84	6.63
4			9.49	
⋮				
10		15.99		

		MARITAL STATUS(MS)			
		MAR.	DIV.	SIN.	
HAIR COLOR (H.C.)	BLOND	125	65	170	360
	BRUN	45	25	50	120
	RED	70	30	140	240
		240	120	360	720

H_0: NO ASSN. BTW H.C. & M.S. H_A: IS. ASSO.

P. EST 1ST R = $360/720$ P. EST 1ST C = $240/720$

P. EST FOR 1ST CELL = $360(240)/720^2$

E 1ST CELL = $nP = \frac{720(360)240}{(720)720} = 120$

MARITAL STATUS(MS)

		MAR.	DIV.	SIN.	
HAIR	BLOND	125 120	65 60	170 180	360
COLOR	BRUN	45 40	25 20	50 60	120
(H.C.)	RED	70 80	30 40	140 120	240
		240	120	360	720

$$\chi^2=\frac{(125-120)^2}{120}+\cdots+\frac{(140-120)^2}{120}=11.805$$

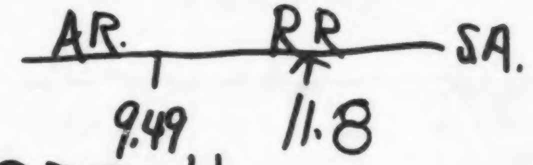
H_0: NO ASSN. BTW H.C.&M.S. H_A: IS. ASSO.

P.EST 1ST R= 360/720 P.EST 1ST C = 240/720

P. EST FOR 1ST CELL= $360(240)/720^2$

$$E\ 1^{ST}\ CELL = mP = \frac{720(360)240}{(720)720} = 120$$

$\alpha = .05$ $\chi^2 = 11.8$



$\chi^2_{.95}(df=4) = 9.49$ REJ H_0

PERCENTILES OF THE χ^2 DISTRIBUTION

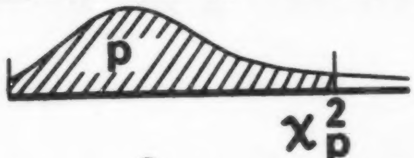


Table A-3
T-24

THE VALUES OF χ^2_p CORRESPONDING TO p

df	...	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.99}$
1			3.84	6.63
4			9.49	
⋮				
10		15.99		

$df > 1$ LARGER THAN 2X2 TA.

AT MOST 20% OF $E < 5$ PROC.

$2 < df < 30$ & MIN. $E \geq 2$ PROCEED

MOST OR ALL $E < 5$ BIOM 1951 P.141

$df > 30$ & MANY SM. E BIOM 1939 P.346

1-DIM. GOODNESS OF FIT
TAIL E MIN AT LEAST 1

CIRCUMSTANCE DECIDED ON A PRIORI GROUNDS
SUSPECT COULD BE IN EITHER TAIL

CASE	CONDITION
I	μ AND σ UNKNOWN ←
II	μ AND σ UNKNOWN EXTERNAL ESTIMATE FOR S AVAILABLE
III	μ UNKNOWN, σ KNOWN
IV	μ AND σ KNOWN

RANKED DATA $\underset{\substack{\uparrow \\ \text{SMALLEST}}}{X_1} \le X_2 \le \cdots \le \underset{\substack{\uparrow \\ \text{LARGEST}}}{X_n}$

H_0: HOMO. SAMPLE, H_A: NOT HOMO.

1, 5, 9, 10, 11, 13, 14, 16, 17, 39 $\alpha = 0.01$

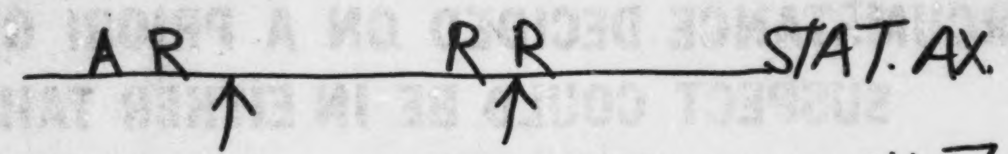
$r_{11} = (X_n - X_{n-1})/(X_n - X_2) = (39-17)/(39-5) = .647$

$C.V = r_{1-0.5\alpha} = r_{0.995} = 0.639$

CRITERIA FOR REJECTION OF OUTLYING OBSERVATIONS

Table A-14
Page T-27

STATISTIC	n	UPPER PERCENTILES		
		.70	. . .	.995
r_{11}	8			
	9			
	10			.639


A R
R R
STAT. AX.

CV=0.639 STATISTIC=0.647

DEC. REJ. H_0 OR CONTAMINATED

SAMPLE OR REJ. 39

$$r_{11} = \frac{X_2 - X_1}{X_{n-1} - X_1}$$ IF X_1 SUSPECTED

CIRCUMSTANCE DECIDED ON A PRIORI GROUNDS SUSPECT COULD BE IN EITHER TAIL

CASE	CONDITION
I	μ AND σ UNKNOWN
II	μ AND σ UNKNOWN ← EXTERNAL ESTIMATE FOR S AVAILABLE
III	μ UNKNOWN, σ KNOWN
IV	μ AND σ KNOWN

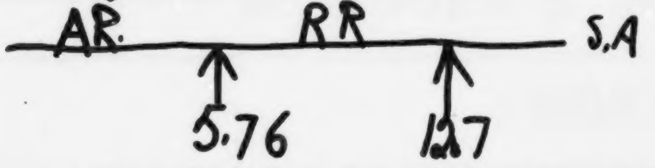
1, 5, 9, 10, 11, 13, 14, 16, 17, 39 $n = 10$

H_0: HOMO. SAMPLE, H_A: NOT HOMO.

$s = 3$ WITH $df = 30 = \nu$ $\alpha = 0.01$

STAT. $= q = (X_n - X_1)/s = (39-1)/3 = 12.7$

C.V. $= q_{1-\alpha}(n, \nu) = q_{0.99}(10, 30) = 5.76$



REJ. H_0

PERCENTILES OF THE STUDENTIZED RANGE, q

Table A-10
T-22

$q_{.99}$

df for s or ν	t = SAMPLE SIZE				
	2	. . .	10	. . .	20
1					
⋮					
30			5.76		
⋮					

CIRCUMSTANCE DECIDED ON A PRIORI GROUNDS SUSPECT IS IN ONE TAIL ONLY

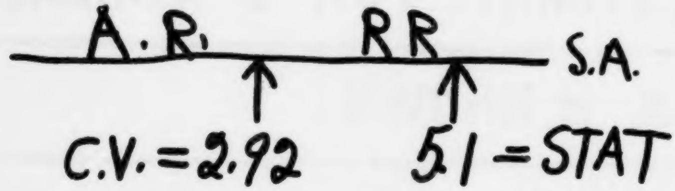
CASE	CONDITION
I	μ AND σ UNKNOWN
II	μ AND σ UNKNOWN EXTERNAL ESTIMATE FOR S AVAILABLE
III	μ UNKNOWN, σ KNOWN ←
IV	μ AND σ KNOWN

$6, 5, \cdots, 17, 39 \quad \Sigma X = 135 \quad \overline{X} = 13.5 \quad \alpha = 0.01$

H_0: HOMO. DATA H_A: NOT HOMO.

$t_m = (X_m - \overline{X})/\sigma = (39 - 13.5)/5 = 5.1$

$C.V. = t_\alpha(n, \nu) = t_{0.01}(10, \infty) \cong 2.88 + .04 = 2.92$



REJ. H_0

PERCENTAGE POINTS OF ... MEAN

Table A-16
T-30

α =0.01 (0.05 ALSO)

df for s or ν	n = SAMPLE SIZE			
	3		9	12
10				
⋮				
120				
∞			2.88	3.01

CIRCUMSTANCE DECIDED ON A PRIORI GROUNDS

SUSPECT IS IN ONE TAIL ONLY

CASE	CONDITION
I	μ AND σ UNKNOWN
II	μ AND σ UNKNOWN EXTERNAL ESTIMATE FOR S AVAILABLE
III	μ UNKNOWN, σ KNOWN
IV	μ AND σ KNOWN ←

$1, 5, \cdots, 17, 39 \quad \mu = 8, \ \sigma = 5, \ \alpha = 0.01$

$0.5\alpha' \equiv 1 - \sqrt[n]{1-\alpha} = 1 - \sqrt[10]{1-0.01} = 0.001$

$Z_{1-0.5\alpha'} = Z_{0.999} = 3.09$

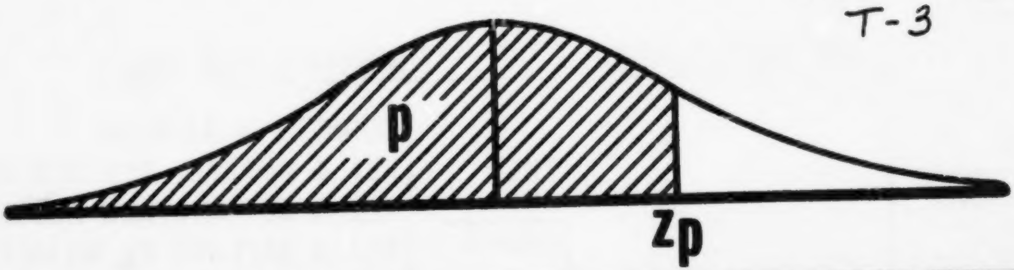
CONF. LIMITS: $\mu \pm \sigma Z = 8 \pm 5(3.09) = -7.5 ; 23.5$

$\underbrace{-7.5 \qquad\qquad 23.5}$ REJ. OBS OUTSIDE INT.

REJ. 39

CUMULATIVE NORMAL DISTRIBUTION — VALUES OF z_p

Table A-2
T-3



P	.9	.95	.975	.999
z_p	1.282	1.645	1.960	3.090

SINGLE OUTLIER IN EITHER TAIL (4)

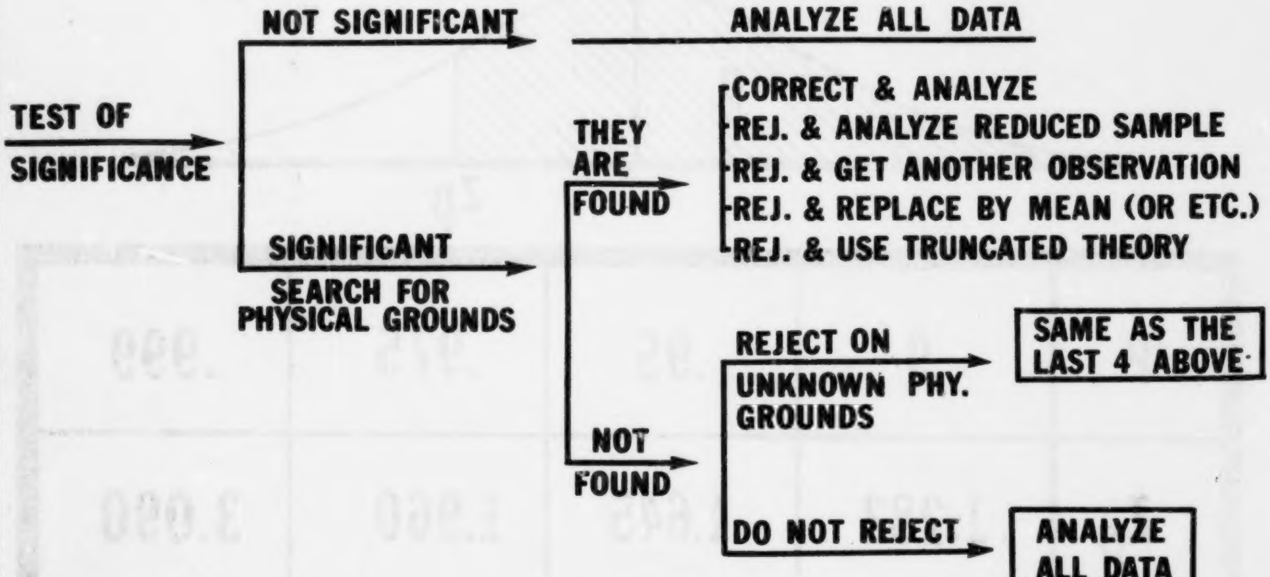
" " " ONE " (4)

" " " BOTH TAILS

TWO OUTLIERS " " "

HIGHLY CONTAMINATED (UP TO 50%)

GRUBBS, TECHNOMETRICS, 1969
VOL. 11 PP 1-21


TEST OF SIGNIFICANCE
NOT SIGNIFICANT
ANALYZE ALL DATA
SIGNIFICANT
SEARCH FOR PHYSICAL GROUNDS
THEY ARE FOUND
CORRECT & ANALYZE
REJ. & ANALYZE REDUCED SAMPLE
REJ. & GET ANOTHER OBSERVATION
REJ. & REPLACE BY MEAN (OR ETC.)
REJ. & USE TRUNCATED THEORY
NOT FOUND
REJECT ON UNKNOWN PHY. GROUNDS
SAME AS THE LAST 4 ABOVE
DO NOT REJECT
ANALYZE ALL DATA

BOTH WAYS